

# Exact simulation of the jump times of a class of Piecewise Deterministic Markov Processes

Vincent Lemaire<sup>\*†</sup>    Michèle Thieullen<sup>\*‡</sup>    Nicolas Thomas<sup>\*§</sup>

## Abstract

In this paper, we are interested in the exact simulation of a class of Piecewise Deterministic Markov Processes (PDMP). We show how to perform efficient thinning algorithms depending on the jump rate bound. For different types of jump rate bounds, we compare theoretically the mean number of generated (total) jump times and we compare numerically the simulation times. We use thinning algorithms on Hodgkin-Huxley models with Markovian ion channels dynamic to illustrate the results.

**Keywords:** Piecewise Deterministic Markov Process; Exact simulation; Thinning; Neural model; Channel noise

**Mathematics Subject Classification:** 60J25, 65C20, 92C20, 92C45

## 1 Introduction

In many areas it is important to be able to simulate exactly and rapidly trajectories of a stochastic process. This is the case for Monte Carlo methods, statistical estimation, bootstrap. In this article, we are interested in the exact simulation (perfect sampling) of a class of Piecewise Deterministic Markov Processes (PDMP). These processes, introduced by M.H.A. Davis in [9], are based on an increasing sequence of random times in which the processes have a jump and on a deterministic evolution between two successive random times. The law of a PDMP is thus determined by three parameters called the characteristics of the PDMP: a family of vector fields, a jump rate (intensity function) and a transition measure. These parameters can be given but in some cases we may have only access to few of them. In this study we suppose that the flows of the PDMP are known, this means that we explicitly know the solution of each ordinary differential equations associated to each vector fields. This hypothesis is not a restriction. Indeed, many random phenomena are modelled by PDMPs with explicit flows. For example, we can quote the temporal evolution of the membrane potential and ionic channels in neuroscience (see [24]), the evolution of a food contaminant in the human body in pharmacokinetic (see [3]), the growth of bacteria in biology (see [12]), the data transmission in internet network (see [5]) or the evolution of a chemical network in chemistry (see [1]).

The flows being known, we focus on the simulation of the jump times. The basic tool we use is

---

<sup>\*</sup>Laboratoire de Probabilités et Modèles Aléatoires, UMR CNRS 7599, UPMC Paris 6, Case 188, 4 place Jussieu, F-75252 Paris Cedex 5, France

<sup>†</sup>vincent.lemaire@upmc.fr

<sup>‡</sup>michele.thieullen@upmc.fr

<sup>§</sup>nicolas.thomas@upmc.fr

the thinning method introduced by Lewis and Schedler in [19] to simulate Poisson processes and generalised by Ogata [22] to simulate any point processes. This method became classic when the jump rate of the process admits a constant upper bound  $\bar{\lambda}$ . It consists in generating the jump times of a (homogeneous) Poisson process with intensity  $\bar{\lambda}$  then selecting some of these times by a rejection argument. In this case, we obtain an algorithm that is easy to implement. However, it is intuitive that a constant bound  $\bar{\lambda}$  will not provide a powerful algorithm when the jump rate of the process present significant variations.

Let us go back to PDMPs. The main result of this article is to develop thinning algorithms where the jump rate bound closely follows the evolution in time of the jump rate of the PDMP to simulate. In the sequel, such a bound is called *optimal bound*.

We see at least three interests in this result. The first is that it applies with weak hypotheses on the jump rate: we are not assuming that the jump rate is bounded, it can depend on time, being monotone or not. The second is that it provides powerful thinning algorithms. The drawback of this method is that more the jump rate bound is close to the jump rate more the method is time consuming. It is thus necessary to look for a satisfactory balance. We discuss this difficulty on a numerical example. Finally, the *optimal bound* is constructed by following each vector fields of the family. This construction is thus natural in the context of switching processes such as PDMPs. For this reason we think that the algorithms studied in this article can be applied to a much larger family of processes such as Hawkes processes or jump (P)SDE with state-dependent intensity.

Let us now give some details on the content of the article. In this study we will consider three types of jump rate bounds:

- the *global bound*, the coarsest, which is constant (in particular it is independent of the state of the PDMP and of time),
- the *local bound*, which depends on each post-jump values of the PDMP and which is constant between two successive jump times,
- the *optimal bound*, the finest, which depends on each post-jump values of the PDMP and also on the time evolution of the process between two successive jump times.

For each of the three jump rate bounds we theoretically compare the mean number of rejected points, then, on a numerical example, the simulation times.

As an indicator of the efficiency of the thinning method, we choose the mean value of the ratio between the number of accepted jump times and the number of proposed jump times. We call it rate of acceptance. This indicator is between 0 and 1 and is easily understood, the closer it is to 1 the less we reject points, thus the more the method is efficient. We explicitly express this rate of acceptance in terms of the transition measure of a discrete time Markov chain which carries more than the information of the PDMP only. Note that it is not the embedded Markov chain of the PDMP since it contains also all the rejected jump times. The rate of acceptance is also expressed as a function of the ratio between the jump rate of the PDMP and the jump rate bound. This expression allows to see that more the jump rate bound is close to the jump rate of the PDMP more the algorithm is efficient. Let us also note that the rate of acceptance is different from the definition of efficiency given in [19] or [11] chap. 6 where it is defined as the ratio between the mean number of accepted jump times and the mean number of proposed jump times. However, in the case of Poisson processes, the rate of acceptance coincides with the definition of efficiency given in [19] or [11] chap. 6.

When the flows are not known explicitly, we can simulate PDMPs with the algorithms given in [25] or in [28] which essentially consist in repeatedly solving ODEs.

We chose to illustrate the results on the efficiency of the *optimal bound* on two stochastic Hodgkin-Huxley (HH) models, the *subunit model* and the *channel model* [24]. The *subunit model* is the stochastic model whose deterministic limit (when the number of channels goes to infinity) is the classical four-dimensional deterministic HH model [16] whereas the deterministic limit of the *channel model* is a fourteen-dimensional model [24]. The difference between the two stochastic models comes from the modelling of the conductance of the membrane. In the *subunit model*, the conductance is given by the proportion of open gates whereas, in the *channel model*, it is given by the proportion of open channels. Note that the two deterministic models are equivalent when the initial conditions satisfy a binomial relation [24]. The *channel model* is known to be more bio-realistic than the *subunit model* but also more complex.

Thinning algorithms are better suited than algorithms in [25] or [28] to simulate variables of biological interest such as mean spiking times or mean dwell times using classical Monte Carlo. Indeed, in addition to be faster, thinning algorithms do not introduce bias since the simulation is exact.

We present both stochastic HH models as PDMPs. The jump rates of these models present high variations especially when the membrane is in a depolarization phase, thus, it allows to check whether the *optimal bound* speed up simulation compared to the *global bound* and the *local bound*. Moreover the jump rates (which come from the modelling [16]) are complex functions, thus numerical inversion of the distribution function can be time consuming. To the best of our knowledge, when flows are known, no studies of the error have been carried out when we numerically inverse the distribution function.

Several algorithms can be found to simulate stochastic HH models. These algorithms can be classified into three categories: channel-state tracking algorithms (CST) [7], [26], [2] channel-number-tracking algorithms (CNT) [27], [6], [2] and approximate algorithms [23], [14], [13]. The main difference between the CNT and CST algorithms comes from the simulation of the inter-jump times and we emphasise that thinning algorithms can be used in both types of algorithm. In the literature, the term exact algorithm (simulation) is employed to denote CST and CNT algorithms (see [13], [2]) even if we use some Euler integration to numerically solve the inversion of the distribution function problem. In this article, the meaning of "exact" signifies that the simulated trajectories are realizations of these processes.

We show how to determine jump rate bounds in such stochastic HH models. We also present a new way to simulate the transition measure of the *channel model* by using the transition measure of the *subunit model*. This approach is numerically more efficient than the classical method used for example in [20] p.587. For the stochastic HH models studied here, we show that the simulation time is reduced by 2 in going from the *global bound* to the *local bound* and that it is also reduced by 2 in going from the *local bound* to the *optimal bound*.

The paper is organised as follows. In section 2, we give the definition of PDMPs and set the notations used in other sections. In section 3, we introduce the jump rate bounds and we present the thinning procedure. In section 4, we give the theoretical results concerning the comparison of the jump rate bounds and the rate of acceptance. In section 5, we introduce the Hodgkin-Huxley models. In section 6 we numerically illustrates the results.

## 2 Formal definition of PDMPs

A PDMP is a stochastic process in which the source of randomness comes from random jump times and post-jump locations. In this paper, we consider that such a process takes the following

general form

$$x_t = (\theta_t, V_t) \quad \forall t \geq 0$$

where:

- $\theta : \mathbb{R}^+ \rightarrow K$  is a jump process that characterizes the mode of the system,  $m \geq 1$ .
- $V : \mathbb{R}^+ \rightarrow D \subseteq \mathbb{R}^d$  is a stochastic process which evolves deterministically between two successive jumps of the process  $\theta$ ,  $d \geq 1$ .
- $K$  is a finite or countable space.

Let us denote  $E = K \times D$  so that  $(x_t)_{t \geq 0}$  is an  $E$ -valued process. We note  $(T_n)_{n \geq 0}$  the sequence of jump times of the PDMP and  $(N_t)_{t \geq 0}$  the counting process  $N_t = \sum_{n \geq 1} \mathbf{1}_{T_n \leq t}$ . We make the following assumption

**Assumption 2.1** : for every starting point  $x \in E$ ,  $\mathbb{E}_x[N_t] < \infty$  for all  $t \geq 0$ .

This assumption implies in particular that  $T_n \rightarrow \infty$  almost surely. Such a PDMP is uniquely defined by its three characteristics  $(\phi, \lambda, Q)$  where

- *The deterministic flow*  $\phi : \mathbb{R}_+ \times E \rightarrow D$  is supposed continuous and induced by a vector field  $F : E \rightarrow D$ . For  $t \in [T_n, T_{n+1}[$ ,  $V$  takes the following form  $V_t = \phi(t - T_n, x_{T_n})$  and the trajectory of the process  $(x_t)_{t \geq 0}$  is then given by

$$x_t = \sum_{n \geq 0} \left( \theta_{T_n}, \phi(t - T_n, x_{T_n}) \right) \mathbf{1}_{T_n \leq t < T_{n+1}}.$$

For notational convenience, we define a vector field  $G : E \rightarrow E$  such that for  $x \in E$   $G(x) = \begin{pmatrix} 0 \\ F(x) \end{pmatrix}$ . Then we can represent the PDMP as follows

$$x_t = \sum_{n \geq 0} \psi(t - T_n, x_{T_n}) \mathbf{1}_{T_n \leq t < T_{n+1}}.$$

Where  $\psi$  is the flow induced by  $G$  such that for  $t \in [T_n, +\infty[$

$$\psi(t - T_n, x_{T_n}) = \left( \theta_{T_n}, \phi(t - T_n, x_{T_n}) \right).$$

- *The jump rate*  $\lambda : E \rightarrow \mathbb{R}_+$  is a non-negative measurable function that characterizes the frequency of jumps and such that for each  $x \in E$  there exists  $\sigma(x)$  such that the function  $s \rightarrow \lambda(\psi(s, x))$  is integrable on  $[0, \sigma(x)[$ . Then the trajectory of the (stochastic) jump rate is given by

$$\lambda(x_t) = \sum_{n \geq 0} \lambda(\psi(t - T_n, x_{T_n})) \mathbf{1}_{T_n \leq t < T_{n+1}}.$$

The formula is to be understood as follows: the intensity of the inter-jump time  $S_{n+1}$  is  $\lambda(\psi(t - T_n, x_{T_n}))$  for  $t \geq T_n$ .

- *The transition measure*  $Q : E \times \mathcal{B}(E) \rightarrow [0, 1]$  governs the post-jump location of the process. It verifies

$$Q(x, \{x\}) = 0 \quad \forall x \in E.$$

The iterative construction of these processes in [9] and [10] provides a way to simulate their trajectories although the problem of the exact simulation of the inter-jump times is not obvious. In [10], M.H.A Davis shows that there exists a filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}_x)$  such that the process  $(x_t)_{t \geq 0}$  is a Markov process. He also shows that  $(x_{T_k})_{k \geq 0}$  is a Markov chain with kernel  $Z$  such that for  $x_0 \in E$

$$\begin{aligned} Z(x_0, A) &= \mathbb{P}(x_{T_1} \in A | x_{T_0} = x_0) \\ &= \int_0^\infty Q(\psi(t, x_0), A) \lambda(\psi(t, x_0)) e^{-\int_0^t \lambda(\psi(s, x_0)) ds} dt. \end{aligned}$$

The randomness of the PDMP is contained in the associated jump process  $(\eta_t)$  defined by

$$\eta_t = x_{T_n} \quad T_n \leq t < T_{n+1}.$$

Because  $T_n = \inf\{t > T_{n-1} : \theta_{t-} \neq \theta_t\}$ , the knowledge of  $(\eta_t)_{t \geq 0}$  implies the knowledge of  $(T_n)_{n \geq 0}$ .

### 3 Simulation of PDMPs by thinning

In this section, we first present the three different jump rate bounds. Secondly, we describe the thinning method to simulate inter-jump times of PDMPs.

#### 3.1 Jump rate bounds

In this section we introduce the different jump rate bounds, namely, the *optimal bound*, the *local bound* and the *global bound*. The *optimal bound* is particularly efficient in term of reject because it is as close as we want to the jump rate. Let us first introduce a general bound of  $\lambda$ , namely  $\tilde{\lambda}$ , defined by

$$\tilde{\lambda}(x_t, t) = \sum_{n \geq 0} \tilde{\lambda}(\psi(t - T_n, x_{T_n}), t) \mathbf{1}_{T_n \leq t < T_{n+1}}$$

where  $(T_n)_{n \geq 0}$  denotes the jump times of the PDMP. For  $n \geq 0$  and conditionally on  $(T_n, x_{T_n})$  the function  $t \mapsto \tilde{\lambda}(\psi(t - T_n, x_{T_n}), t)$  is defined on  $[T_n, +\infty[$  and verifies

$$\lambda(\psi(t - T_n, x_{T_n})) \leq \tilde{\lambda}(\psi(t - T_n, x_{T_n}), t) \quad \forall t \geq T_n.$$

We shall precise that the function of time  $\tilde{\lambda}(\psi(t - T_n, x_{T_n}), t)$  is used to simulate  $S_{n+1}$ . For  $n \geq 0$  and conditionally on  $(T_n, x_{T_n})$  we define the function  $\tilde{\Lambda}_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  by

$$\tilde{\Lambda}_n(t) = \int_{T_n}^t \tilde{\lambda}(\psi(s - T_n, x_{T_n}), s) ds$$

We denote by  $(\tilde{\Lambda}_n)^{-1}$  the inverse of  $\tilde{\Lambda}_n$ .

##### 3.1.1 The global bound

We define the *global bound* by

$$\tilde{\lambda}^{\text{glo}}(x_t, t) = \sum_{n \geq 0} \sup_{x \in E} \lambda(x) \mathbf{1}_{T_n \leq t < T_{n+1}}$$

By definition, this bound is constant and does not depend on the state of the PDMP nor on time, we will denote it by  $\tilde{\lambda}^{\text{glo}}$ . This bound is probably the most used and has the advantage to lead to an easy implementation. Indeed, to simulate the jump times of the PDMP we simulate a homogeneous Poisson process with jump rate  $\tilde{\lambda}^{\text{glo}}$  disregarding the state of the PDMP. For  $n \geq 0$  and conditionally on  $(T_n, x_{T_n})$ , the integrated jump rate bound is given by  $\tilde{\Lambda}_n^{\text{glo}}(t) = \tilde{\lambda}^{\text{glo}}(t - T_n)$  for  $t \geq T_n$  and the inverse is given by  $(\tilde{\Lambda}_n^{\text{glo}})^{-1}(s) = s/\tilde{\lambda}^{\text{glo}} + T_n$  for  $s \geq 0$ .

### 3.1.2 The local bound

We define the *local bound* by

$$\tilde{\lambda}^{\text{loc}}(x_t, t) \equiv \sum_{n \geq 0} \left( \sup_{t \geq T_n} \lambda(\psi(t - T_n, x_{T_n})) \right) \mathbf{1}_{T_n \leq t < T_{n+1}}$$

By definition, this bound is constant between two successive jump times and has the advantage of being adapted to the state of the PDMP right-after a jump. To each jump time of the PDMP corresponds a homogeneous Poisson process whose intensity depends on the state of the PDMP at the jump time. For  $n \geq 0$  and conditionally on  $(T_n, x_{T_n})$ , the integrated jump rate bound is  $\tilde{\Lambda}_n^{\text{loc}}(t) = \left( \sup_{t \geq T_n} \lambda(\psi(t - T_n, x_{T_n})) \right) (t - T_n)$  for  $t \geq T_n$  and the inverse is given by  $(\tilde{\Lambda}_n^{\text{loc}})^{-1}(s) = \left( s / \sup_{t \geq T_n} \lambda(\psi(t - T_n, x_{T_n})) \right) + T_n$  for  $s \geq 0$ .

### 3.1.3 The optimal bound

Let  $P$  be a finite or a countable space. For  $n \geq 0$ , let us denote by  $(\mathcal{P}_k^{T_n})_{k \in P}$  a partition of  $[T_n, +\infty[$ . Thus, for  $k \in P$ ,  $\mathcal{P}_k^{T_n}$  is an interval of  $[T_n, +\infty[$ . We define the *optimal bound* by

$$\tilde{\lambda}^{\text{opt}}(x_t, t) \equiv \sum_{n \geq 0} \left( \sum_{k \in P} \sup_{t \in \mathcal{P}_k^{T_n}} \lambda(\psi(t - T_n, x_{T_n})) \mathbf{1}_{\mathcal{P}_k^{T_n}}(t) \right) \mathbf{1}_{T_n \leq t < T_{n+1}}$$

By definition, this bound is piecewise constant between two successive jump times, thus it is adapted to the state of the PDMP right-after a jump but also to the evolution in time of the jump rate. To each jump time of the PDMP corresponds a non-homogeneous Poisson process whose intensity depends on the state of the PDMP at the jump time and on the flow starting from this state. For  $n \geq 0$  and conditionally on  $(T_n, x_{T_n})$ , the integrated jump rate bound is, for  $t \geq T_n$ ,

$$\tilde{\Lambda}_n^{\text{opt}}(t) = \sum_{k \in P} \sup_{t \in \mathcal{P}_k^{T_n}} \lambda(\psi(t - T_n, x_{T_n})) \left| \mathcal{P}_k^{T_n} \cap [T_n, t] \right|$$

where  $\left| \mathcal{P}_k^{T_n} \cap [T_n, t] \right|$  represents the length of the interval  $\mathcal{P}_k^{T_n} \cap [T_n, t]$ . As  $P$  is at most countable, let us denote by  $p_i$  for  $i = 0, \dots, \text{Card}(P)$  its elements. The inverse,  $(\tilde{\Lambda}_n^{\text{opt}})^{-1}$ , is given by

$$(\tilde{\Lambda}_n^{\text{opt}})^{-1}(s) = \sum_{i=0}^{\text{card}(P)} \left( \frac{s - \sum_{k=0}^{i-1} \sup_{t \in \mathcal{P}_{p_k}^{T_n}} \lambda(\psi(t - T_n, x_{T_n})) \left| \mathcal{P}_{p_k}^{T_n} \right|}{\sup_{t \in \mathcal{P}_{p_i}^{T_n}} \lambda(\psi(t - T_n, x_{T_n}))} + T_n + \sum_{l=0}^{i-1} \left| \mathcal{P}_{p_l}^{T_n} \right| \right) \mathbf{1}_{[\kappa_{p_{i-1}}, \kappa_{p_i}]}(s)$$

where  $\kappa_{p_i} = \sum_{k=0}^i \sup_{t \in \mathcal{P}_{p_k}^{T_n}} \lambda(\psi(t - T_n, x_{T_n})) \left| \mathcal{P}_{p_k}^{T_n} \right|$  and, by convention  $\sum_{l=0}^{-1} \left| \mathcal{P}_{p_l}^{T_n} \right| = 0$  and  $\kappa_{p_{-1}} = 0$ .

**Remark 1:** Let  $P = \mathbb{N}$  and let  $\epsilon > 0$ . The *optimal bound* with the partition  $(\mathcal{P}_k^{T_n, \epsilon})_{k \in \mathbb{N}}$  where, for  $k \in \mathbb{N}$ ,  $\mathcal{P}_k^{T_n, \epsilon} = [T_n + k\epsilon, T_n + (k+1)\epsilon[$  can be use even if the jump rate  $\lambda$  is not bounded. A bounded hypothesis on  $\lambda$  is required to use the *global bound* and it can be weakened for the *local bound* (for example if the flows are bounded and  $\lambda$  continuous).

**Remark 2:** For the three jump rate bounds, the simulation is exact. In particular, for all finite or countable  $P$ , that is, for any partitions of  $[T_n, +\infty[$ , the simulation remains exact.

**Remark 3:** The choice of the bound depends on the PDMP we want to simulate. If the jump rate does not vary very much in time, the *local bound* or the *global constant* bound can be adopted but if the jump rate presents high variations in a small time interval, the *optimal bound* is preferable.

### 3.2 Thinning

In this section, we detail the thinning procedure to simulate a PDMP. Details on the thinning of Poisson processes may be find in [19] or [11]. We present the procedure with the general bound  $\tilde{\lambda}$  of section 3.1. In practice, one has to apply the procedure with one of the three bounds (*optimal*, *local* or *global*). We simulate a sample path of the PDMP  $(x_t)_{t \geq 0}$  with values in  $E$ , starting from a fixed initial point  $x_{T_0} = x_0$  at time  $T_0$  as follows.

Let  $(\tilde{T}_k^0)_{k \geq 0}$  be the Poisson process defined on  $[0, +\infty[$  with jump rate  $\tilde{\lambda}(\psi(t, x_0), t)$  for  $t \geq 0$ , and

$$\tau_1 = \inf\{k > 0 : U_k \tilde{\lambda}(\psi(\tilde{T}_k^0, x_0), \tilde{T}_k^0) \leq \lambda(\psi(\tilde{T}_k^0, x_0))\}$$

where  $(U_n)_{n \geq 1}$  is a sequence of independent random variables with uniform distribution on  $[0, 1]$ , independent of  $(\tilde{T}_k^0)_{k \geq 0}$  and  $x_0$ . The first jump time  $T_1 = S_1$  of the PDMP is the first jump of a non-homogeneous Poisson process defined on  $[0, +\infty[$  with jump rate  $\lambda(\psi(t, x_0))$  for  $t \geq 0$ . We simulate  $T_1$  by thinning the process  $(\tilde{T}_k^0)_{k \geq 0}$ , then,  $T_1 = \tilde{T}_{\tau_1}^0$ . On  $[0, T_1[$  the PDMP evolves as follows

$$x_t = \psi(t, x_0)$$

and the random variable  $x_{T_1}$  has distribution  $Q((\psi(T_1, x_0)), \cdot)$ . Note that conditionally on  $(T_0, x_{T_0}, T_1)$  the process  $(\tilde{T}_k^0)_{k \geq 1}$  is a Poisson process on  $[T_0, T_1[$  with jump rate  $\tilde{\lambda}(\psi(t, x_0), t) - \lambda(\psi(t, x_0))$ , see [11] chap.6.

Suppose we have simulated  $T_i$ , then, conditionally on  $(T_i, x_{T_i})$ , the PDMP  $(x_t)$  restarts from  $x_{T_i}$  at time  $T_i$  independently from the past. Let  $(T_i + \tilde{T}_k^i)_{k \geq 0}$  be the Poisson process on  $[T_i, +\infty[$  with jump rate  $\tilde{\lambda}(\psi(t - T_i, x_{T_i}), t)$  for  $t \geq T_i$  and

$$\tau_{i+1} = \inf\{k > 0 : U_k \tilde{\lambda}(\psi(\tilde{T}_k^i, x_{T_i}), T_i + \tilde{T}_k^i) \leq \lambda(\psi(\tilde{T}_k^i, x_{T_i}))\}$$

where  $(U_n)_{n \geq 1}$  is a sequence of independent uniform random variables, independent of  $(\tilde{T}_k^i)_{k \geq 0}$  and  $x_{T_i}$ .  $T_{i+1}$  is the first jump of a non-homogeneous Poisson process defined on  $[T_i, +\infty[$  with jump rate  $\lambda(\psi(t - T_i, x_{T_i}))$  for  $t \geq T_i$ . We simulate  $T_{i+1}$  by thinning the process  $(T_i + \tilde{T}_k^i)_{k \geq 0}$ , then,  $T_{i+1} = T_i + \tilde{T}_{\tau_{i+1}}^i$ . On  $[T_i, T_{i+1}[$  the process evolves as follows

$$x_t = \psi(t - T_i, x_{T_i})$$

and the random variable  $x_{T_{i+1}}$  has distribution  $Q((\psi(S_{i+1}, x_{T_i})), \cdot)$ . Note that, conditionally on  $(T_i, x_{T_i}, T_{i+1})$ , the process  $(T_i + \tilde{T}_k^i)_{k \geq 1}$  is a Poisson process on  $[T_i, T_{i+1}[$  with jump rate  $\tilde{\lambda}(\psi(t - T_i, x_{T_i}), t) - \lambda(\psi(t - T_i, x_{T_i}))$ .

Moreover, conditionally in  $(T_0, x_{T_0}, \dots, T_i, x_{T_i}, T_{i+1})$ , the points in  $[T_i, T_{i+1}[$  obtained from the

Poisson process  $(T_i + \tilde{T}_k^i)_{k \geq 1}$  are independent of the points in  $[T_{j-1}, T_j[$  obtained from the Poisson process  $(T_j + \tilde{T}_k^j)_{k \geq 1}$  for  $j = 1, \dots, i$ . Thus, we have constructed a point process

$$T_0 < \tilde{T}_1^0 < \dots < \tilde{T}_{\tau_1-1}^0 < T_1 < T_1 + \tilde{T}_1^1 < \dots < T_1 + \tilde{T}_{\tau_2-1}^1 < T_2 < T_2 + \tilde{T}_1^2 < \dots \quad (1)$$

**Notation:** In the sequel, the process defined by (1) is noted  $(\tilde{T}_k)_{k \geq 0}$  and the associated counting process is noted  $(\tilde{N}_t)_{t \geq 0}$ . We also denote by  $(\bar{T}_k)_{k \geq 0}$  the process formed by all the rejected points (i.e the process  $(\tilde{T}_k)_{k \geq 0}$  without the jump times  $(T_k)_{k \geq 0}$ ) and  $(\bar{N}_t)_{t \geq 0}$  the associated counting process.

The sequence  $(\tilde{T}_k)_{k \geq 0}$  contains both rejected and accepted points and the sub-sequence noted  $(T_k)_{k \geq 0}$  such that for  $k \geq 1$ ,  $T_k = \sum_{l=1}^k \tilde{T}_{\tau_l}^{l-1}$  defines the jump times of the PDMP. Thus, we have constructed the jumps of the PDMP by thinning the process  $(\tilde{T}_k)_{k \geq 0}$  with non-constant probabilities  $(p_k)$  such that  $p_k = \lambda(x_{\tilde{T}_k-}) / \tilde{\lambda}(x_{\tilde{T}_k-}, \tilde{T}_k)$  is the probability to accept  $\tilde{T}_k$ .

An important fact in the procedure is that  $(\tilde{T}_k)_{k \geq 0}$  is composed by pieces of independent Poisson processes  $(\tilde{T}_k^0), (\tilde{T}_k^1), \dots, (\tilde{T}_k^i), \dots$

The thinning procedure provides an algorithm to simulate trajectories of the PDMP.

### Algorithm

---

#### Step 1.

Fix the initial time  $T_0$ , the initial condition  $x_{T_0} = (\theta_{T_0}, V_{T_0})$  and set a jump counter  $n = 0$ . Set also an auxiliary jump counter  $k = 0$  and an auxiliary variable  $\tilde{T}_k = T_n$ .

#### Step 2.

$k \leftarrow k + 1$

Simulate  $U_{2k-1} \sim \mathcal{U}(]0, 1[)$

Simulate  $E_k = -\log(U_{2k-1})$

$\tilde{T}_k = (\tilde{\Lambda}_n)^{-1}(E_k + \tilde{\Lambda}_n(\tilde{T}_{k-1}))$

#### Step 3.

Simulate  $U_{2k} \sim \mathcal{U}(]0, 1[)$

if  $U_{2k} \tilde{\lambda}(\psi(\tilde{T}_k - T_n, x_{T_n}), \tilde{T}_k) > \lambda(\psi(\tilde{T}_k - T_n, x_{T_n}))$  go to **Step 2**

#### Step 4.

$T_{n+1} = \tilde{T}_k$

Let  $V_t = \phi(t - T_n, x_{T_n})$  for  $t \in [T_n, T_{n+1}[$ .

If  $T_{n+1} > T$  stop at  $t = T$ .

#### Step 5.

Otherwise, simulate a post-jump value  $x_{T_{n+1}}$  according to the Markovian kernel  $Q(\psi(S_{n+1}, x_{T_n}), \cdot)$ .

#### Step 6.

Set  $n = n + 1$  and return to **Step 2**.

---

## 4 Efficiency of the thinning procedure

In this section, we compare the efficiency of the thinning method in term of reject for the different bounds. Recall that the number of points needed to simulate one inter-jump time of the PDMP is given by

$$\tau_i^x = \inf\{k > 0 : U_k \tilde{\lambda}^x(\psi(\tilde{T}_k^{i-1}, x_{T_{i-1}}), T_{i-1} + \tilde{T}_k^{i-1}) \leq \lambda(\psi(\tilde{T}_k^{i-1}, x_{T_{i-1}}))\}$$



for  $i \geq 1$ ,  $x \in \{\text{glo}, \text{loc}, \text{opt}\}$  where  $(U_n)_{n \geq 1}$  and  $(\tilde{T}_k^{i-1})_{k \geq 0}$  are defined as in section 3.2. We begin by a lemma.

**Lemma 4.1:** We have the following uniform convergence

$$\sup_{x \in E} \sup_{t \geq 0} |\tilde{\lambda}^\epsilon(t, x) - \lambda(\psi(t, x))| \xrightarrow{\epsilon \rightarrow 0} 0$$

where

$$\tilde{\lambda}^\epsilon(t, x) = \sum_{k \geq 0} \sup_{s \in [k\epsilon, (k+1)\epsilon[} \lambda(\psi(s, x)) \mathbf{1}_{[k\epsilon, (k+1)\epsilon[}(t)$$

*Proof.* For  $n > 0$  we set  $\epsilon = 1/n$ , thus

$$\tilde{\lambda}^{1/n}(t, x) = \sum_{k \geq 0} \sup_{s \in [k/n, (k+1)/n[} \lambda(\psi(s, x)) \mathbf{1}_{[k/n, (k+1)/n[}(t).$$

Let  $M = \sup_{x \in E} \sup_{t \geq 0} \left| \frac{\partial \lambda}{\partial t}(\psi(t, x)) \right|$ ,  $\nu > 0$ ,  $N = \lceil M/\nu \rceil$  and  $n \geq N$ . Let  $x \in E$  and  $t \geq 0$ , there exists  $l \geq 0$  such that  $t \in [l/n, (l+1)/n[$ . Thus

$$\tilde{\lambda}^{1/n}(t, x) = \sup_{s \in [l/n, (l+1)/n[} \lambda(\psi(s, x)).$$

Let  $t_0 \in [l/n, (l+1)/n]$  such that  $\sup_{s \in [l/n, (l+1)/n[} \lambda(\psi(s, x)) = \lambda(\psi(t_0, x))$ . The application of the mean value inequality to the function  $\lambda \circ \psi$  gives

$$|\lambda(\psi(t_0, x)) - \lambda(\psi(t, x))| \leq M|t_0 - t| \leq M \frac{1}{n} \leq \nu.$$

And the conclusion follows.  $\square$

## 4.1 Comparison of the mean number of total jump times

In this section, the  $\tau_i^x$  for  $x \in \{\text{glo}, \text{loc}, \text{opt}\}$  are called local reject. In proposition 4.1, we show that the best local reject is obtained with the *optimal bound*. More the local reject is small less pseudo-random variables have to be simulated. Thus, the simulation time using the *optimal bound* is expected to be smaller than with the two other bounds.

**Proposition 4.1:** For all  $i \geq 1$ , We have

$$\mathbb{E}[\tau_i^{\text{opt}} | x_{T_{i-1}}, T_{i-1}] \leq \mathbb{E}[\tau_i^{\text{loc}} | x_{T_{i-1}}, T_{i-1}] \leq \mathbb{E}[\tau_i^{\text{glo}} | x_{T_{i-1}}, T_{i-1}].$$

*Proof.* Let  $i \geq 1$  and  $S_i = T_i - T_{i-1}$ . We apply theorem 2.2 of [11] chap.6 and obtain

$$\begin{aligned} \mathbb{E}[\tau_i^{\text{glo}} | x_{T_{i-1}}, T_{i-1}] &= \tilde{\lambda}^{\text{glo}} \int_{T_{i-1}}^{+\infty} e^{-\int_{T_{i-1}}^t \lambda(\psi(s-T_{i-1}, x_{T_{i-1}})) ds} dt = \tilde{\lambda}^{\text{glo}} \mathbb{E}[S_i | x_{T_{i-1}}, T_{i-1}] \\ \mathbb{E}[\tau_i^{\text{loc}} | x_{T_{i-1}}, T_{i-1}] &= \sup_{t \geq T_{i-1}} \lambda(\psi(t - T_{i-1}, x_{T_{i-1}})) \int_{T_{i-1}}^{+\infty} e^{-\int_{T_{i-1}}^t \lambda(\psi(s-T_{i-1}, x_{T_{i-1}})) ds} dt \\ &= \sup_{t \geq T_{i-1}} \lambda(\psi(t - T_{i-1}, x_{T_{i-1}})) \mathbb{E}[S_i | x_{T_i}, T_{i-1}] \\ \mathbb{E}[\tau_i^{\text{opt}} | x_{T_{i-1}}, T_{i-1}] &= \sum_{k \geq 0} \sup_{s \in \mathcal{P}_k^{T_{i-1}}} \lambda(\psi(s - T_{i-1}, x_{T_{i-1}})) \int_{\mathcal{P}_k^{T_{i-1}}} e^{-\int_{T_{i-1}}^t \lambda(\psi(s-T_{i-1}, x_{T_{i-1}})) ds} dt. \end{aligned}$$

We have

$$\tilde{\lambda}^{\text{glo}} = \sup_{x \in E} \lambda(x) \geq \sup_{t \geq T_{i-1}} \lambda(\psi(t - T_{i-1}, x_{T_{i-1}})) \geq \sup_{t \in \mathcal{P}_k^{T_{i-1}}} \lambda(\psi(t - T_{i-1}, x_{T_{i-1}}))$$

for all  $k \geq 0$ . Since, for all  $i \geq 1$ ,  $\mathbb{E}[S_i | x_{T_{i-1}}, T_{i-1}] \geq 0$ , the conclusion follows.  $\square$

From Proposition 4.1, we deduce that  $\mathbb{E}[\tilde{N}_t^{\text{opt}}] \leq \mathbb{E}[\tilde{N}_t^{\text{loc}}] \leq \mathbb{E}[\tilde{N}_t^{\text{glo}}]$  where  $\tilde{N}_t^{\text{opt}}$ ,  $\tilde{N}_t^{\text{loc}}$  and  $\tilde{N}_t^{\text{glo}}$  are counting processes with intensity  $\tilde{\lambda}_t^{\text{opt}}$ ,  $\tilde{\lambda}_t^{\text{loc}}$  and  $\tilde{\lambda}_t^{\text{glo}}$  respectively.

## 4.2 Rate of acceptance

We are now interested in the rate of acceptance, that is the mean proportion of accepted points in an interval of the form  $[0, t]$  for  $t > 0$ . Let  $(N_t)$  be the counting process of the PDMP and  $(\tilde{N}_t)$  the counting process with general jump rate  $\tilde{\lambda}$  (see section 3.1). In proposition 4.2 we give an explicit formula of the rate of acceptance defined as  $\mathbb{E}[N_t / \tilde{N}_t | \tilde{N}_t \geq 1]$ . Recall that, for  $k \geq 1$ ,

$$p_k = \lambda(\psi(\tilde{T}_k - T_{n_k}, x_{T_{n_k}})) / \tilde{\lambda}(\psi(\tilde{T}_k - T_{n_k}, x_{T_{n_k}}), \tilde{T}_k)$$

is the probability to accept the point  $\tilde{T}_k$  where  $T_{n_k}$  denotes the last accepted point before  $\tilde{T}_k$  and  $(\tilde{T}_k)$  is defined by (1). Let  $J : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be the process defined by  $J_t = \sum_{k \geq 0} (t - T_k) \mathbf{1}_{T_k \leq t < T_{k+1}}$ , thus, for  $t \geq 0$ ,  $J_t$  gives the age of the last accepted point before  $t$ . Then, for  $k \geq 1$ , we can write the probabilities  $p_k$  as follows

$$p_k = \lambda(\psi(J_{\tilde{T}_{k-1}} + \tilde{S}_k, \eta_{\tilde{T}_{k-1}})) / \tilde{\lambda}(\psi(J_{\tilde{T}_{k-1}} + \tilde{S}_k, \eta_{\tilde{T}_{k-1}}), \tilde{T}_k)$$

where  $\tilde{S}_k = \tilde{T}_k - \tilde{T}_{k-1}$  and  $(\eta_t)$  is as in section 2. The process  $(\tilde{S}_k, \tilde{T}_k, J_{\tilde{T}_k}, \eta_{\tilde{T}_k})_{k \geq 0}$  defines a Markov chain on  $\mathbb{R}_+ \times \tilde{E}$  where  $\tilde{E} = \mathbb{R}_+ \times \mathbb{R}_+ \times E$  with semi-Markovian kernel

$$M(\tilde{T}_{k-1}, J_{\tilde{T}_{k-1}}, \eta_{\tilde{T}_{k-1}}, ds, dt, dj, dx) = \tilde{Q}(s, \tilde{T}_{k-1}, J_{\tilde{T}_{k-1}}, \eta_{\tilde{T}_{k-1}}, dt, dj, dx) \alpha(\tilde{T}_{k-1}, J_{\tilde{T}_{k-1}}, \eta_{\tilde{T}_{k-1}}, ds)$$

where

$$\alpha(\tilde{T}_{k-1}, J_{\tilde{T}_{k-1}}, \eta_{\tilde{T}_{k-1}}, ds) = \tilde{\lambda}(\psi(J_{\tilde{T}_{k-1}} + s, \eta_{\tilde{T}_{k-1}}), \tilde{T}_{k-1} + s) e^{-\int_0^s \tilde{\lambda}(\psi(J_{\tilde{T}_{k-1}} + z, \eta_{\tilde{T}_{k-1}}), \tilde{T}_{k-1} + z) dz} ds$$

and

$$\begin{aligned} \tilde{Q}(s, \tilde{T}_{k-1}, J_{\tilde{T}_{k-1}}, \eta_{\tilde{T}_{k-1}}, dt, dj, dx) = \\ \left(1 - \frac{\lambda(\psi(J_{\tilde{T}_{k-1}} + s, \eta_{\tilde{T}_{k-1}}))}{\tilde{\lambda}(\psi(J_{\tilde{T}_{k-1}} + s, \eta_{\tilde{T}_{k-1}}), t)}\right) \delta_{J_{\tilde{T}_{k-1}} + s}(dj) \delta_{\eta_{\tilde{T}_{k-1}}}(dx) \delta_{\tilde{T}_{k-1} + s}(dt) + \\ \frac{\lambda(\psi(J_{\tilde{T}_{k-1}} + s, \eta_{\tilde{T}_{k-1}}))}{\tilde{\lambda}(\psi(J_{\tilde{T}_{k-1}} + s, \eta_{\tilde{T}_{k-1}}), t)} Q\left(\lambda(\psi(J_{\tilde{T}_{k-1}} + s, \eta_{\tilde{T}_{k-1}})), dx\right) \delta_0(dj) \delta_{\tilde{T}_{k-1} + s}(dt) \end{aligned}$$

**Proposition 4.2:** Let  $(N_t)_{t \geq 0}$  be the counting process of the PDMP  $(x_t)_{t \geq 0}$ ,  $(\tilde{N}_t)_{t \geq 0}$  be the counting process with jump times  $(\tilde{T}_n)_{n \geq 0}$  and  $M$  be the kernel of the Markov chain  $(\tilde{S}_k, \tilde{T}_k, J_{\tilde{T}_k}, \eta_{\tilde{T}_k})_{k \geq 0}$ , we have

$$\begin{aligned} \mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] = \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \int_{\mathbb{R}_+} \int_{\tilde{E}} \sum_{k=1}^n \left( \frac{\lambda(\psi(j_{k-1} + s_k, x_{k-1}))}{\tilde{\lambda}(\psi(j_{k-1} + s_k, x_{k-1}), t_k)} \right) e^{-\int_0^{t-t_n} \tilde{\lambda}(\psi(j_n + z, x_n), t_n + z) dz} \\ \mathbf{1}_{t \geq t_n} \mu(dx_0) M(0, 0, x_0, ds_1, dt_1, dj_1, dx_1) \dots M(t_{n-1}, j_{n-1}, x_{n-1}, ds_n, dt_n, dj_n, dx_n) \end{aligned}$$

where  $\mu$  is the law of  $\eta_{\tilde{T}_0}$ .

*Proof.* We provide a proof into two steps. First, we establish that, with an appropriate conditioning, the conditional law of  $N_t$  is the conditional law of a sum of independent Bernoulli random variables with different parameters. Then, we use this property as well as the kernel  $M$  to compute the rate of acceptance.

Let  $n \geq 1$  and let us define  $n$  independent Bernoulli random variables  $X_i$  with parameters  $p_i$  such that

$$p_i = \frac{\lambda(\psi(J_{\tilde{T}_{i-1}} + \tilde{S}_i, \eta_{\tilde{T}_{i-1}}))}{\tilde{\lambda}(\psi(J_{\tilde{T}_{i-1}} + \tilde{S}_i, \eta_{\tilde{T}_{i-1}}), \tilde{T}_i)}$$

We note  $X = \sum_{i=1}^n X_i$  the sum of the Bernoulli variables and  $A_{t,n} = \{\tilde{N}_t = n, p_1, \dots, p_n\}$ . By noting that, for  $0 \leq k \leq n$ , we have

$$\{N_t = k | A_{t,n}\} = \bigcup_{1 \leq i_1 < \dots < i_k \leq n} \left[ \bigcap_{i \in I_k^n} \{U_i \leq p_i\} \bigcap_{i \in \bar{I}_k^n} \{U_i > p_i\} \right] = \{X = k | p_1, \dots, p_n\}$$

where  $I_k^n = \{i_1, \dots, i_k\} \subseteq \{1, \dots, n\}$ ,  $\bar{I}_k^n$  the complementary of  $I_k^n$  in  $\{1, \dots, n\}$  and  $(U_i)$  are independent random variables uniformly distributed in  $[0, 1]$  and independent of  $(p_i)$ , we deduce that

$$\mathcal{L}(N_t | A_{t,n}) = \mathcal{L}(X | p_1, \dots, p_n).$$

In particular,  $\mathbb{E}[N_t | A_{t,n}] = \mathbb{E}[X | p_1, \dots, p_n] = \sum_{i=1}^n p_i$ . Thus, one can write

$$\begin{aligned} \mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] &= \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}[N_t \mathbf{1}_{\tilde{N}_t = n}] \\ &= \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}[\mathbb{E}[N_t | A_{t,n}] | \tilde{N}_t = n] \mathbb{P}(\tilde{N}_t = n) \\ &= \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n p_i \mathbf{1}_{\tilde{N}_t = n}\right] \\ &= \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n p_i \mathbf{1}_{t - \tilde{T}_n \geq 0} \mathbb{E}\left[\mathbf{1}_{\tilde{S}_{n+1} \geq t - \tilde{T}_n} | \eta_{\tilde{T}_0}, \tilde{S}_1, \dots, \tilde{S}_n, \tilde{T}_n, J_{\tilde{T}_n}, \eta_{\tilde{T}_n}\right]\right] \end{aligned}$$

Conditionally to  $(\eta_{\tilde{T}_0}, \tilde{S}_1, \tilde{T}_1, J_{\tilde{T}_1}, \eta_{\tilde{T}_1}, \dots, \tilde{S}_n, \tilde{T}_n, J_{\tilde{T}_n}, \eta_{\tilde{T}_n})$ , the random variable  $\tilde{S}_{n+1}$  is a hazard law with rate  $\tilde{\lambda}(\psi(J_{\tilde{T}_n} + t, \eta_{\tilde{T}_n}), \tilde{T}_n + t)$  for  $t \geq 0$ , thus

$$\begin{aligned} \mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] &= \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n p_i e^{-\int_0^{t - \tilde{T}_n} \tilde{\lambda}(\psi(J_{\tilde{T}_n} + u, \eta_{\tilde{T}_n}), \tilde{T}_n + u) du} \mathbf{1}_{t - \tilde{T}_n \geq 0}\right] \\ &= \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}\left[f(\eta_{\tilde{T}_0}, \tilde{S}_1, \tilde{T}_1, J_{\tilde{T}_1}, \eta_{\tilde{T}_1}, \dots, \tilde{S}_n, \tilde{T}_n, J_{\tilde{T}_n}, \eta_{\tilde{T}_n})\right] \end{aligned}$$

Where

$$\begin{aligned} f(\eta_{\tilde{T}_0}, \tilde{S}_1, \tilde{T}_1, J_{\tilde{T}_1}, \eta_{\tilde{T}_1}, \dots, \tilde{S}_n, \tilde{T}_n, J_{\tilde{T}_n}, \eta_{\tilde{T}_n}) &= \\ &= \sum_{i=1}^n \left( \frac{\lambda(\psi(J_{\tilde{T}_{i-1}} + \tilde{S}_i, \eta_{\tilde{T}_{i-1}}))}{\tilde{\lambda}(\psi(J_{\tilde{T}_{i-1}} + \tilde{S}_i, \eta_{\tilde{T}_{i-1}}), \tilde{T}_i)} \right) e^{-\int_0^{t - \tilde{T}_n} \tilde{\lambda}(\psi(J_{\tilde{T}_n} + u, \eta_{\tilde{T}_n}), \tilde{T}_n + u) du} \mathbf{1}_{t - \tilde{T}_n \geq 0} \end{aligned}$$

Since  $(\tilde{S}_k, \tilde{T}_k, J_{\tilde{T}_k}, \eta_{\tilde{T}_k})_{k \geq 0}$  is a Markov chain with kernel  $M$ , we obtain

$$\mathbb{E}[f(\eta_{\tilde{T}_0}, \tilde{S}_1, \dots, \tilde{S}_n, J_{\tilde{T}_n}, \eta_{\tilde{T}_n})] = \int_{\mathbb{R}_+} \int_{\tilde{E}} f(x_0, s_1, t_1, j_1, x_1, \dots, s_n, t_n, j_n, x_n) \\ \mu(dx_0) M(0, 0, x_0, ds_1, dt_1, dj_1, dx_1) \dots M(t_{n-1}, j_{n-1}, x_{n-1}, ds_n, dt_n, dj_n, dx_n)$$

Where  $\mu$  is the law of  $\eta_{\tilde{T}_0}$ . Thus we have the result.  $\square$

When  $\tilde{\lambda}$  is close to  $\lambda$ , the rate of acceptance is expected to be close to 1. As an example, consider the case of two Poisson processes  $(N_t)$  and  $(\tilde{N}_t)$  with intensity  $\lambda(t)$  and  $\tilde{\lambda}(t)$  respectively such that  $\tilde{\lambda}(t) = \tilde{\lambda}$  for all  $t \geq 0$ . Thus, for  $n \geq 1$ ,  $\tilde{S}_1, \dots, \tilde{S}_n$  are independent exponential variables with parameter  $\tilde{\lambda}$ . Let us also consider that  $\lambda(t) \simeq \tilde{\lambda}$  for  $t \geq 0$ . In this case, the rate of acceptance is

$$\mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] \simeq \frac{1}{1 - e^{-\tilde{\lambda}t}} \sum_{n \geq 1} \int_{(\mathbb{R}_+)^n} e^{-\tilde{\lambda}(t - (s_1 + \dots + s_n))} \mathbf{1}_{t \geq s_1 + \dots + s_n} \alpha(ds_1), \dots, \alpha(ds_n)$$

where  $\alpha(ds) = \tilde{\lambda} e^{-\tilde{\lambda}s} ds$ . Since,  $\tilde{T}_n = \tilde{S}_1 + \dots + \tilde{S}_n$  is gamma distributed with parameters  $n$  and  $\tilde{\lambda}$ , we have

$$\mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] \simeq \frac{1}{1 - e^{-\tilde{\lambda}t}} \sum_{n \geq 1} \mathbb{E}[e^{-\tilde{\lambda}(t - \tilde{T}_n)} \mathbf{1}_{t \geq \tilde{T}_n}] \simeq \frac{1}{1 - e^{-\tilde{\lambda}t}} \sum_{n \geq 1} \frac{(\tilde{\lambda}t)^n}{n!} e^{-\tilde{\lambda}t} \simeq 1$$

### 4.3 Convergence of the counting process with a specific optimal bound as jump rate

Let  $P = \mathbb{N}$  and let  $\epsilon > 0$ , we consider the optimal bound with the partition  $(\mathcal{P}_k^{T_n, \epsilon})_{k \in \mathbb{N}}$  where, for  $k \in \mathbb{N}$ ,  $\mathcal{P}_k^{T_n, \epsilon} = [T_n + k\epsilon, T_n + (k+1)\epsilon[$ , we denote it by  $\tilde{\lambda}^{\text{opt}, \epsilon}(x_t, t)$  and we note  $(\tilde{N}_t^{\text{opt}, \epsilon})$  the corresponding counting process. The number of points needed to simulate one inter-jump time with this particular partition is noted  $\tau_i^{\text{opt}, \epsilon}$ .

We first show, in proposition 4.3, that  $(\tilde{T}_k)_{k \geq 0}$  (defined in section 3.2) is a Cox process with stochastic jump rate  $\tilde{\lambda}(x_t, t) - \lambda(x_t)$ . Details on Cox processes can be found in [18].

Then we show, in proposition 4.4, that the counting process  $(\tilde{N}_t^\epsilon)$  whose jump rate is  $\tilde{\lambda}^{\text{opt}, \epsilon}$  converge to the counting process  $(N_t)$  of the PDMP.

Proposition 4.5 states that more the parameter  $\epsilon$  is small less rejected points are simulated. However, when  $\epsilon$  is too small, step 2 of the algorithm requires many iterations to compute  $\tilde{\Lambda}(\cdot)$  and  $(\tilde{\Lambda})^{-1}(\cdot)$  and the simulation time increases. We will see in the numerical section 6.2.1 that taking an  $\epsilon$  of order  $\max_n(T_{n+1} - T_n)$  leads to the optimal simulation time.

**Proposition 4.3** : Let  $\xi$  be a point process and  $\mu$  be a random measure such that

$$\xi([0, t]) = \sum_{n \geq 0} \mathbf{1}_{\tilde{T}_n \leq t} \quad \text{and} \quad \mu([0, t]) = \int_0^t \tilde{\lambda}(x_s, s) - \lambda(x_s) ds$$

Then  $\xi$  is a Cox process directed by  $\mu$ .

*Proof.* Let us first note that for  $t \geq 0$

$$\tilde{\lambda}(x_t, t) - \lambda(x_t) = \sum_{n \geq 0} \left[ \tilde{\lambda}(\psi(t - T_n, x_{T_n}), t) - \lambda(\psi(t - T_n, x_{T_n})) \right] \mathbf{1}_{T_n \leq t < T_{n+1}}$$

Thanks to [18], we show that for any measurable and non-negative functions  $f$ , the Laplace functional of  $\xi$  is

$$\mathbb{E}[e^{-\xi f}] = \mathbb{E}[e^{-\mu(1-e^{-f})}]$$

Let  $f$  be a non-negative measurable function. Let us note  $f_T(t) = f(t)\mathbf{1}_{t \leq T}$  for  $T > 0$ , such that  $\lim_{T \rightarrow \infty} f_T(t) = f(t)$  with  $f_T$  increasing with  $T$ . Thus, by Beppo-Levi's theorem,  $\xi f_T \nearrow \xi f$  and then  $e^{-\xi f_T} \searrow e^{-\xi f}$  when  $T$  goes to infinity. Moreover,  $e^{-\xi f_T} \leq 1$ , thus by Lebesgue's Dominated Convergence Theorem

$$\mathbb{E}[e^{-\xi f_T}] \rightarrow \mathbb{E}[e^{-\xi f}]$$

With the same type of arguments, we show that

$$\mathbb{E}[e^{-\mu(1-e^{-f_T})}] \rightarrow \mathbb{E}[e^{-\mu(1-e^{-f})}]$$

Thus, it is sufficient to show (2) for functions  $f_T$ .

$$\begin{aligned} \mathbb{E}[e^{-\xi f_T}] &= \mathbb{E}[e^{-\sum_{n \geq 1} f_T(\bar{T}_n)}] \\ &= \sum_{k \geq 0} \mathbb{E}[e^{-\sum_{n \geq 1} f_T(\bar{T}_n)} | N_T = k] \mathbb{P}(N_T = k) \\ &= \sum_{k \geq 0} \mathbb{E}\left[\mathbb{E}\left[\prod_{i=0}^k e^{-\sum_{n \geq 1} f_T(\bar{T}_n) \mathbf{1}_{T_i \leq \bar{T}_n < T_{i+1}}} | N_T = k, (\eta_t)_{0 \leq t \leq T} | N_T = k\right] \mathbb{P}(N_T = k)\right] \end{aligned}$$

By the thinning procedure, the points  $\bar{T}_n$  in  $[T_i, T_{i+1}[$  may be written as  $T_i + \tilde{T}_l^i$  for some  $l \geq 1$  where  $(\tilde{T}_l^i)_{l \geq 1}$  is, conditionally on  $x_{T_i}$ , a Poisson process with jump rate  $\tilde{\lambda}(\psi(t - T_i, x_{T_i}), t) - \lambda(\psi(t - T_i, x_{T_i}))$  for  $t \geq T_i$ . Since  $(\tilde{T}_l^i)$  is independent of  $(\tilde{T}_l^j)$  for  $i \neq j$ , the random variables  $X_i := e^{-\sum_{n \geq 1} f_T(\bar{T}_n) \mathbf{1}_{T_i \leq \bar{T}_n < T_{i+1}}}$  are independent. Thus, since the Laplace functional of a Poisson process  $\xi$  with intensity  $\mu$  verifies  $\mathbb{E}[e^{-\xi f}] = e^{-\mu(1-e^{-f})}$ , we obtain

$$\begin{aligned} \mathbb{E}[e^{-\xi f_T}] &= \sum_{k \geq 0} \mathbb{E}\left[\prod_{i=0}^k \mathbb{E}[e^{-\sum_{n \geq 1} f_T(\bar{T}_n) \mathbf{1}_{T_i \leq \bar{T}_n < T_{i+1}}} | N_T = k, (\eta_t)_{0 \leq t \leq T} | N_T = k] \mathbb{P}(N_T = k)\right] \\ &= \sum_{k \geq 0} \mathbb{E}\left[e^{-\sum_{i=0}^{N_T} \int \left(1 - e^{-f_T(s) \mathbf{1}_{T_i \leq s < T_{i+1}}}\right) \left(\tilde{\lambda}(\psi(s - T_i, x_{T_i}), s) - \lambda(\psi(s - T_i, x_{T_i}))\right) ds} | N_T = k\right] \mathbb{P}(N_T = k) \\ &= \mathbb{E}\left[e^{-\sum_{i \geq 0} \int \left(1 - e^{-f_T(s)}\right) \mathbf{1}_{T_i \leq s < T_{i+1}} \left(\tilde{\lambda}(\psi(s - T_i, x_{T_i}), s) - \lambda(\psi(s - T_i, x_{T_i}))\right) ds}\right] \\ &= \mathbb{E}[e^{-\mu(1-e^{-f_T})}] \end{aligned}$$

□

**Proposition 4.4:** Let  $(\tilde{N}_t^\epsilon)$  be the counting process whose jump rate is  $\tilde{\lambda}^{\text{opt}, \epsilon}$  and let  $(N_t)$  be the counting process of the PDMP  $x_t$ , we have the following convergence in law

$$\tilde{N}^\epsilon \xrightarrow{\epsilon \rightarrow 0} N$$

*Proof.* Let  $f$  be a non-negative measurable function. We show the convergence of the Laplace transform [8], that is

$$\mathbb{E}[e^{-\int f d\tilde{N}^\epsilon}] \xrightarrow{\epsilon \rightarrow 0} \mathbb{E}[e^{-\int f dN}].$$

Let  $T > 0$ , as in proposition 4.3, it is sufficient to show the convergence of the Laplace transform for functions  $f_T(t) = f(t)\mathbf{1}_{t \leq T}$ . Let  $(\tilde{T}_n^\epsilon)$  be the points of the process  $\tilde{N}^\epsilon$ .

$$\begin{aligned} \mathbb{E}[e^{-\int f_T d\tilde{N}^\epsilon}] &= \mathbb{E}[e^{-\sum_{n \geq 0} f_T(\tilde{T}_n^\epsilon)}] \\ &= \mathbb{E}\left[\mathbb{E}[e^{-\sum_{n \geq 0} f_T(\tilde{T}_n^\epsilon)} | (\eta_t)_{0 \leq t \leq T}]\right] \\ &= \mathbb{E}\left[e^{-\sum_{n \geq 0} f_T(T_n)} \mathbb{E}[e^{-\sum_{n \geq 0} f_T(\bar{T}_n^\epsilon)} | (\eta_t)_{0 \leq t \leq T}]\right] \end{aligned}$$

Where  $(\bar{T}_n^\epsilon)$  denotes the rejected points. Since  $(\bar{T}_n^\epsilon)$  is a Cox process with stochastic jump rate  $\tilde{\lambda}^{\text{opt}, \epsilon}(x_t, t) - \lambda(x_t)$ , we obtain

$$\mathbb{E}[e^{-\int f_T d\tilde{N}^\epsilon}] = \mathbb{E}\left[e^{-\sum_{n \geq 0} f_T(T_n)} e^{-\int (1 - e^{-f_T(s)}) (\tilde{\lambda}^{\text{opt}, \epsilon}(x_s, s) - \lambda(x_s)) ds}\right]$$

Since  $e^{-\sum_{n \geq 0} f_T(T_n)} e^{-\int (1 - e^{-f_T(s)}) (\tilde{\lambda}^{\text{opt}, \epsilon}(x_s, s) - \lambda(x_s)) ds} \leq 1$ , we obtain by Lebesgue's Dominated Convergence Theorem and by continuity of the exponential

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}[e^{-\int f_T d\tilde{N}^\epsilon}] = \mathbb{E}\left[e^{-\sum_{n \geq 0} f_T(T_n)} e^{-\lim_{\epsilon \rightarrow 0} \int (1 - e^{-f_T(s)}) (\tilde{\lambda}^{\text{opt}, \epsilon}(x_s, s) - \lambda(x_s)) ds}\right].$$

We have

$$-T \sup_{x \in E} \sup_{t \geq 0} \left( \tilde{\lambda}^\epsilon(t, x) - \lambda(\psi(t, x)) \right) \leq \int -(1 - e^{-f_T(s)}) (\tilde{\lambda}^{\text{opt}, \epsilon}(x_s, s) - \lambda(x_s)) ds \leq 0$$

where

$$\tilde{\lambda}^\epsilon(t, x) = \sum_{k \geq 0} \sup_{s \in [k\epsilon, (k+1)\epsilon[} \lambda(\psi(s, x)) \mathbf{1}_{[k\epsilon, (k+1)\epsilon[}(t)$$

By lemma 4.1, we obtain that almost surely  $e^{\lim_{\epsilon \rightarrow 0} \int (1 - e^{-f_T(s)}) (\tilde{\lambda}^{\text{opt}, \epsilon}(x_s, s) - \lambda(x_s)) ds} = 1$ , the conclusion follows since  $\mathbb{E}\left[e^{-\sum_{n \geq 0} f_T(T_n)}\right] = \mathbb{E}\left[e^{-\int f_T dN}\right]$ .  $\square$

**Proposition 4.5:** For all  $i \geq 1$  we have

$$\mathbb{E}[\tau_i^{\text{opt}, \epsilon} | x_{T_{i-1}}, T_{i-1}] \xrightarrow{\epsilon \rightarrow 0} 1$$

*Proof.* Let  $i \geq 1$  and  $\epsilon > 0$ . From theorem 2.2 in chap.6 of [11] we have

$$\mathbb{E}[\tau_i^{\text{opt}, \epsilon} | x_{T_{i-1}}, T_{i-1}] = \int_{T_{i-1}}^{+\infty} \tilde{\lambda}_{i-1}^\epsilon(t - T_{i-1}) e^{-\int_{T_{i-1}}^t \lambda(\psi(s - T_{i-1}, x_{T_{i-1}})) ds} dt$$

Where

$$\tilde{\lambda}_{i-1}^\epsilon(t - T_{i-1}) = \sum_{k \geq 0} \sup_{s \in \mathcal{P}_k^{T_{i-1}}} \lambda(\psi(s - T_{i-1}, x_{T_{i-1}})) \mathbf{1}_{\mathcal{P}_k^{T_{i-1}, \epsilon}}(t)$$

from theorem 2.3 in [11] chap.6, we deduce that

$$\mathbb{E}[\tau_i^{\text{opt},\epsilon}|x_{T_{i-1}}, T_{i-1}] \leq \sup_{t \geq T_{i-1}} \frac{\tilde{\lambda}_{i-1}^\epsilon(t - T_{i-1})}{\lambda(\psi(t - T_{i-1}, x_{T_{i-1}}))}.$$

By Lemma 4.1, we obtain

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}[\tau_i^{\text{opt},\epsilon}|x_{T_{i-1}}, T_{i-1}] \leq 1.$$

Since  $\mathbb{E}[\tau_i^{\text{opt},\epsilon}|x_{T_{i-1}}, T_{i-1}] \geq 1$  for all  $\epsilon > 0$ , the conclusion follows.  $\square$

## 5 Hodgkin-Huxley models

In this section we apply the thinning algorithms above to different stochastic versions of the Hodgkin-Huxley model. First, we introduce the classical deterministic model and the stochastic models, then we give essential points of the simulation.

### 5.1 Deterministic Hodgkin-Huxley model

Alan Lloyd Hodgkin and Andrew Huxley provided two models in 1952 to explain the ionic mechanisms underlying the initiation and propagation of action potentials in the squid giant axon [16].

In the present section, we only consider models of initiation of action potentials. It means that we clamp (isolate) a piece of the axon or of the soma and we study electrical properties in time in this clamped area (also called membrane-patch).

Neuron's membrane separates the intracellular environment from the extracellular one and allows exchanges of material and energy between these two environments. These exchanges are allowed by the opening and closing of gates located on the membrane. In most neurons, the intracellular environment contains a large proportion of potassium ions, whereas the extracellular environment contains a majority of sodium ones. Hodgkin and Huxley discovered that the initiation of action potentials principally relies on the movement of these two kind of ions across the membrane via ionic channels. A ionic channel is constituted by four gates which can be of different types (activation and inactivation) and is specific to one type of ions, for example, a sodium channel allows sodium ions only to pass the membrane. We say that a channel is active when all his gates are open.

A stimulation (it can be an input from other neurons or external applied current) makes the sodium channels active, thus sodium ions enter in the intracellular environment : the membrane is depolarized. This depolarizing step increases the conductance of the membrane and when the voltage exceeds a threshold value there is an action potential. After being active, sodium channels become inactive, while potassium gates open (these opening make the potassium channels active). Potassium ions leave the intracellular environment to compensate the entry of sodium ions : the membrane is re-polarized. Potassium channels stay active longer than sodium ones : the membrane is hyper-polarized. Then, a protein makes the sodium ions go back into the intracellular environment and expels potassium ions outside. These are the principal steps of the initiation of an action potential.

In the models studied in this section, we distinguish three types of channels : sodium, potassium, and leak. We consider that a sodium (potassium) channel is composed by three activation gates represented by the variable  $m$  and one inactivation gate  $h$  (four activation gates  $n$  and zero inactivation gates) and that leak channels are always open and allow all type of ions to pass the membrane.

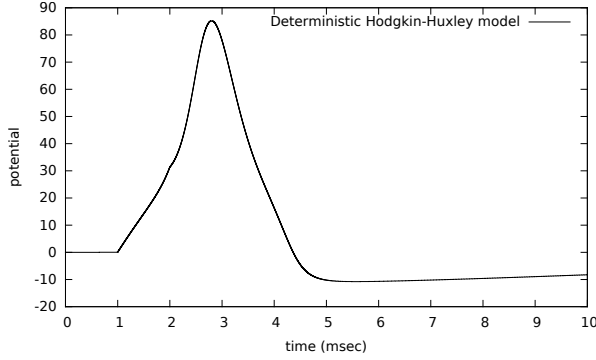


Figure 1: A trajectory of the deterministic Hodgkin-Huxley model with the values of the parameters given in appendix A and  $I(t) = 30\mathbf{1}_{[1,2]}(t)$ .

In fact, there exists other ionic channels as for example calcium ones. In some cases, the calcium plays the role of the sodium to initiate action potentials as in the giant barnacle muscle fiber [21], or more generally in crustacean muscles [15]. Calcium channels have also many other functionalities [15]. Classically, the Hodgkin-Huxley model is the set of nonlinear differential equations

$$\begin{aligned} C \frac{dV}{dt} &= I - I_L(V) - I_{Na}(V, m, h) - I_K(V, n) \\ \frac{dm}{dt} &= (1 - m)\alpha_m(V) - m\beta_m(V) \\ \frac{dh}{dt} &= (1 - h)\alpha_h(V) - h\beta_h(V) \\ \frac{dn}{dt} &= (1 - n)\alpha_n(V) - n\beta_n(V) \end{aligned}$$

The function  $V$  represents the membrane potential (voltage). The functions  $m, h, n$  correspond to the fraction of open gates of type  $m, h$ , or  $n$ . The functions  $\alpha_x$  and  $\beta_x$  for  $x = m, h, n$  are opening and closing rates of gates  $x$  respectively.  $I$  is a time-dependent function which represents the input current,  $C$  is the membrane capacity. For  $z \in \{\text{Na}, \text{K}, \text{L}\}$ ,  $I_z = \bar{g}_z(V - V_z)$  represents the ionic currents where  $\bar{g}_{Na} = g_{Na}m^3h$ ,  $\bar{g}_K = g_Kn^4$  and  $\bar{g}_L = g_L$  are the conductances of the sodium, potassium and leak respectively.  $g_L, g_{Na}, g_K$  are the conductances when all gates are opened and  $V_L, V_{Na}, V_K$  are the resting potentials. Note that, the conductance of the membrane depends on the state of the gates. Thus, more the open gates are, more the conductance is high. This model describes the electrical behaviour of a neuron with an infinite number of gates. Thus, it does not reflect the variability observed experimentally.

## 5.2 Stochastic Hodgkin-Huxley models

Neurons are subject to various sources of fluctuations, intrinsic (from the membrane) and extrinsic (from synapses). The intrinsic fluctuations are mainly caused by ion channels. To take into account these fluctuations in the model, we fix a finite number of sodium channels  $N_{Na}$  and potassium ones  $N_K$ , and replace the deterministic dynamic of the gates by stochastic processes. Here, we discuss about two stochastic models, namely the *subunit model* and the *channel model*. These models belong to the class of Piecewise Deterministic Markov Processes.



### 5.2.1 The subunit model

The *subunit model* is obtained by considering that the conductance of the membrane depends on the empirical measure defined by the proportion of open gates. Recall that  $m$  (respectively  $n$ ) denotes an activation gate of the sodium (respectively potassium) channels and  $h$  an inactivation one. We define the number of gates  $m$  (respectively  $h, n$ ) by  $N_m = 3N_{\text{Na}}$  (respectively  $N_h = N_{\text{Na}}$ ,  $N_n = 4N_{\text{K}}$ ). Let us consider that each gate is represented by a  $\{0, 1\}$ -valued Markovian Jump Process (MJP) noted  $u_k^{(x)}$  for  $x = m, h, n$  and  $k = 1, \dots, N_x$ . State 1 corresponds to the open configuration and 0 to the close one. Recall also that the opening and closing rates which depend on the voltage are noted  $\alpha_x(\cdot)$  and  $\beta_x(\cdot)$  respectively. The dynamic of a gate can be represented by the following diagram

$$\begin{array}{ccc} & \alpha_x(\cdot) & \\ & \xrightarrow{\quad} & \\ 0 & & 1 \\ & \xleftarrow{\quad} & \\ & \beta_x(\cdot) & \end{array}$$

We consider that all MJPs are independent and we define the number of open gates  $x$  at time  $t$  by

$$\theta^{(x)}(t) = \sum_{k=1}^{N_x} u_k^{(x)}(t).$$

Furthermore, let  $\Theta_{\text{sub}} = \{0, \dots, N_n\} \times \{0, \dots, N_m\} \times \{0, \dots, N_h\}$  be the state space of the process  $\theta_t = (\theta^{(n)}(t), \theta^{(m)}(t), \theta^{(h)}(t))$  which records the number of open gates at time  $t$ . Note that,  $N_x - \theta^{(x)}(t)$  gives the number of close gates  $x$  at time  $t$ . The *subunit model* takes the following form

$$(S) \quad \begin{cases} C \frac{dV_t}{dt} = f^{\text{sub}}(\theta_t, V_t, t) \\ (\theta_t) \end{cases}$$

Where

$$f^{\text{sub}}(\theta, V, t) = I(t) - g_L(V - V_L) - g_{\text{Na}} N_m^{-3} (\theta^{(m)})^3 N_h^{-1} \theta^{(h)} (V - V_{\text{Na}}) - g_{\text{K}} N_n^{-4} (\theta^{(n)})^4 (V - V_{\text{K}})$$

We also define the jump rate of the process by

$$\begin{aligned} \lambda^{\text{sub}}(\theta, V) = & \left( \alpha_m(V)(N_m - \theta^{(m)}) + \beta_m(V)\theta^{(m)} \right) + \left( \alpha_h(V)(N_h - \theta^{(h)}) + \beta_h(V)\theta^{(h)} \right) + \\ & \left( \alpha_n(V)(N_n - \theta^{(n)}) + \beta_n(V)\theta^{(n)} \right) \end{aligned}$$

The component  $V$  represents the membrane potential and is continuous, thus, the Markovian kernel  $Q^{\text{sub}}$  is only concerned by the post-jump location of the jump process  $\theta$ . We suppose that two channels do not change states simultaneously almost surely. For example, the probability of the event of exactly one gate  $n$  opens (conditionally on the last jump time being  $T_k$ ) is given by

$$Q^{\text{sub}}\left((\theta_{T_{k-1}}, V_{T_k}), \{\theta_{T_{k-1}} + (1, 0, 0)\}\right) = \frac{\alpha_n(V_{T_k})(N_n - \theta^{(n)}(T_{k-1}))}{\lambda^{\text{sub}}(\theta_{T_{k-1}}, V_{T_k})}$$

To summarize, the *subunit model* can be expressed as a PDMP  $x_t^{\text{sub}} = (\theta_t, V_t, t) \in \Theta_{\text{sub}} \times \mathbb{R} \times \mathbb{R}_+$  with vector field  $f^{\text{sub}} : \Theta_{\text{sub}} \times \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , jump rate  $\lambda^{\text{sub}} : \Theta_{\text{sub}} \times \mathbb{R} \rightarrow \mathbb{R}_+$ , and a Markovian kernel  $Q^{\text{sub}} : \Theta_{\text{sub}} \times \mathbb{R} \times \mathcal{B}(\Theta_{\text{sub}}) \rightarrow [0, 1]$ .

### 5.2.2 The channel model

In the *channel model*, we form groups of four gates to make channels. Unlike the *subunit model*, we define independent MJPs  $u_k^{(\text{Na})}$  for  $k = 1, \dots, N_{\text{Na}}$  (respectively  $u_k^{(\text{K})}$  for  $k = 1, \dots, N_{\text{K}}$ ) to model the sodium (respectively potassium) channels. These independent MJPs follow kinetic scheme given in appendix B. Note that the conducting state (the state that makes the channel active) of sodium (respectively potassium) channels is  $\{m_3h_1\}$  (respectively  $\{n_4\}$ ) which corresponds to three open gates  $m$  and one open gate  $h$  (respectively four open gates  $n$ ). The conductance of the membrane depends on the empirical measure defined by the proportion of active channels and we define the number of active channels at time  $t \geq 0$  by

$$\begin{aligned}\theta^{(m_3h_1)}(t) &= \sum_{k=1}^{N_{\text{Na}}} \mathbf{1}_{\{m_3h_1\}} \left( u_k^{(\text{Na})}(t) \right), \\ \theta^{(n_4)}(t) &= \sum_{k=1}^{N_{\text{K}}} \mathbf{1}_{\{n_4\}} \left( u_k^{(\text{K})}(t) \right).\end{aligned}$$

For  $i = 0, 1, 2, 3$  and  $j = 0, 1$ , let  $\theta^{(m_ih_j)}$  be the number of channels in state  $\{m_ih_j\}$  and for  $k = 0, 1, 2, 3, 4$ , let  $\theta^{(n_k)}$  be the number of channels in state  $\{n_k\}$ . Let

$$\Theta_{\text{chan}} = \{\theta \in \{0, \dots, N_{\text{Na}}\}^8 \times \{0, \dots, N_{\text{K}}\}^5 : \sum_{i=0}^3 \sum_{j=0}^1 \theta^{(m_ih_j)} = N_{\text{Na}}, \sum_{k=0}^5 \theta^{(n_k)} = N_{\text{K}}\}$$

be the state space of the process  $\theta_t = \left( (\theta^{(m_ih_j)}(t))_{i,j}, (\theta^{(n_k)}(t))_k \right)$  which records the configuration of the channels at time  $t$ . The *channel model* takes the following form

$$(C) \quad \begin{cases} C \frac{dV_t}{dt} = f^{\text{chan}}(\theta_t, V_t, t) \\ (\theta_t) \end{cases}$$

where

$$f^{\text{chan}}(\theta, V, t) = I(t) - g_{\text{L}}(V - V_{\text{L}}) - g_{\text{Na}} N_{\text{Na}}^{-1} \theta^{(m_3h_1)} (V - V_{\text{Na}}) - g_{\text{K}} N_{\text{K}}^{-1} \theta^{(n_4)} (V - V_{\text{K}}).$$

A change in the configuration of the channels (which can be observable or not unlike the *subunit model* in which all changes are observable) happens when a gate opens or closes. We define the application  $\eta : \Theta_{\text{chan}} \rightarrow \Theta_{\text{sub}}$  which, given a configuration of channels, returns the configuration of the corresponding gates. We have

$$\eta(\theta) = \begin{bmatrix} \theta^{(n_1)} + 2\theta^{(n_2)} + 3\theta^{(n_3)} + 4\theta^{(n_4)} \\ \theta^{(m_1h_0)} + 2\theta^{(m_2h_0)} + 3\theta^{(m_3h_0)} + \theta^{(m_1h_1)} + 2\theta^{(m_2h_1)} + 3\theta^{(m_3h_1)} \\ \theta^{(m_0h_1)} + \theta^{(m_1h_1)} + \theta^{(m_2h_1)} + \theta^{(m_3h_1)} \end{bmatrix}$$

The first component of the vector  $\eta(\theta)$  contains  $\theta_{\text{open}}^n$ , the number of open gates  $n$ , the second  $\theta_{\text{open}}^m$ , the number of open gates  $m$  and the third  $\theta_{\text{open}}^h$ , the number of open gates  $h$ . Thus, for  $x = m, h, n$ ,  $\theta_{\text{close}}^x(t) = N_x - \theta_{\text{open}}^x(t)$  gives the number of close gates  $x$  at time  $t$ . We define the jump rate of the *channel model* by

$$\lambda^{\text{chan}}(\theta, V) = \lambda^{\text{sub}}(\eta(\theta), V)$$

Where

$$\lambda^{\text{sub}}(\eta(\theta), V) = \left( \alpha_m(V)(N_m - \theta_{\text{open}}^m) + \beta_m(V)\theta_{\text{open}}^m \right) + \left( \alpha_h(V)(N_h - \theta_{\text{open}}^h) + \beta_h(V)\theta_{\text{open}}^h \right) + \left( \alpha_n(V)(N_n - \theta_{\text{open}}^n) + \beta_n(V)\theta_{\text{open}}^n \right)$$

Since  $V$  is continuous, the kernel  $Q^{\text{chan}}$  is also only concerned by the post-location of the process  $\theta$ . Define  $Q^{\text{chan}}$  as it is classically do in the literature ([25] p.53 and [20] p.587) is computationally expensive because we have more transitions to deal with than in the *subunit model*. We propose to decompose the kernel  $Q^{\text{chan}}$  into a product of two kernels. The decomposition is based on the following observation : it is a change in the configuration of the gates that implies a change in the configuration of the channels. Thus, to determine which transition occurs at time  $t$  among the 28 transitions given in appendix B, we first determine which gate opens or closes by using the kernel  $Q^{\text{sub}}$  with  $\lambda^{\text{sub}}(\eta(\cdot), \cdot)$  and then, depending on which gate changes state, we determine a channel transition by using another kernel. For example, suppose that at time  $t$  a gate  $m$  opens, thus, the possible channel transitions are :  $\{m_0h_0 \rightarrow m_1h_0\}$ ,  $\{m_1h_0 \rightarrow m_2h_0\}$ ,  $\{m_2h_0 \rightarrow m_3h_0\}$ ,  $\{m_0h_1 \rightarrow m_1h_1\}$ ,  $\{m_1h_1 \rightarrow m_2h_1\}$ ,  $\{m_2h_1 \rightarrow m_3h_1\}$  and the next transition is one of those. We define six kernels to take into account all the possibilities.

Let  $L_{\text{open}}^m$ ,  $L_{\text{close}}^m$ ,  $L_{\text{open}}^h$ ,  $L_{\text{close}}^h$ ,  $L_{\text{open}}^n$ ,  $L_{\text{close}}^n$  be kernels defined on  $\Theta_{\text{chan}} \times \mathbb{R} \times \mathcal{B}(\Theta_{\text{chan}})$  with values in  $[0, 1]$  such that  $L_{\text{open}}^m$  is the kernel which choose a transition as above,  $L_{\text{open}}^h$  is a kernel which choose a transition among the following ones  $\{m_0h_0 \rightarrow m_0h_1\}$ ,  $\{m_1h_0 \rightarrow m_1h_1\}$ ,  $\{m_2h_0 \rightarrow m_2h_1\}$ ,  $\{m_3h_0 \rightarrow m_3h_1\}$  and so on. For example, the probability of the event of having the transition  $\{m_0h_0 \rightarrow m_1h_0\}$  (conditional on the last jump time being  $T_k$ ) is given by

$$Q^{\text{chan}}\left((\theta_{T_{k-1}}, V_{T_k}), \{\theta_{T_{k-1}} + (-1, +1, 0, \dots, 0)\}\right) = Q^{\text{sub}}\left((\eta(\theta_{T_{k-1}}), V_{T_k}), \{\eta(\theta_{T_{k-1}}) + (0, 1, 0)\}\right) \times L_{\text{open}}^m\left((\theta_{T_{k-1}}, V_{T_k}), \{\theta_{T_{k-1}} + (-1, +1, 0, \dots, 0)\}\right)$$

Where

$$Q^{\text{sub}}\left((\eta(\theta_{T_{k-1}}), V_{T_k}), \{\eta(\theta_{T_{k-1}}) + (0, 1, 0)\}\right) = \frac{\alpha_m(V_{T_k})\theta_{\text{close}}^m(T_{k-1})}{\lambda^{\text{sub}}(\eta(\theta_{T_{k-1}}), V_{T_k})}$$

$$L_{\text{open}}^m\left((\theta_{T_{k-1}}, V_{T_k}), \{\theta_{T_{k-1}} + (-1, +1, 0, \dots, 0)\}\right) = \frac{3\theta^{(m_0h_0)}(T_{k-1})}{\theta_{\text{close}}^m(T_{k-1})}$$

Finally, the probability of having the transition  $\{m_0h_0 \rightarrow m_1h_0\}$  is, as expected, given by the rate of this transition multiplied by the number of channels in the state  $\{m_0h_0\}$  divided by the total rate.

For  $x \in E$ , the support  $K_x^{\text{chan}}$  of the discrete measure of probability  $Q^{\text{chan}}(x, \cdot)$  contains at most 28 elements (depending on the current state  $x$ ), thus, in the worst case we have to do 28 "if - then" tests to determine the next transition. With the decomposition of  $Q^{\text{chan}}$ , we have, in the worst case 12 "if - then" tests to do. Indeed, for  $x \in E$  the support  $K_x^{\text{sub}}$  of the discrete probability  $Q^{\text{sub}}(\eta(x), \cdot)$  contains at most six elements, and the support of the probabilities  $L_{\text{open}}^m(x, \cdot)$ ,  $L_{\text{close}}^m(x, \cdot)$ ,  $L_{\text{open}}^h(x, \cdot)$ ,  $L_{\text{close}}^h(x, \cdot)$ ,  $L_{\text{open}}^n(x, \cdot)$ ,  $L_{\text{close}}^n(x, \cdot)$  contains also at most six elements (when we deal with a transition of a gate  $m$ ). Therefore, it is computationally cheaper to decompose the kernel.

Thus, the *channel model* can be expressed as a PDMP  $x_t^{\text{chan}} = (\theta_t, V_t, t) \in \Theta_{\text{chan}} \times \mathbb{R} \times \mathbb{R}_+$  with vector field  $f^{\text{chan}} : \Theta_{\text{chan}} \times \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , jump rate  $\lambda^{\text{chan}} : \Theta_{\text{chan}} \times \mathbb{R} \rightarrow \mathbb{R}_+$ , and a Markovian kernel  $Q^{\text{chan}} : \Theta_{\text{chan}} \times \mathbb{R} \times \mathcal{B}(\Theta_{\text{chan}}) \rightarrow [0, 1]$ .

### 5.2.3 Explicit form of the flow of the PDMP between two successive jump times

In this section, we determine the explicit expression of the flow of both models. For  $n \geq 0$ ,  $t \geq T_n$  and  $x \in \{\text{sub}, \text{chan}\}$ , the trajectory of the flow  $\phi$  on  $[T_n, +\infty[$  is given by the following ODE

$$\begin{cases} \frac{d\phi(t-T_n, x_{T_n})}{dt} = f^x\left(\theta_{T_n}, \phi(t-T_n, x_{T_n}), t\right) = -a_n^x \phi(t-T_n, x_{T_n}) + b_n^x + \frac{1}{C} I(t) \\ \phi(0, x_{T_n}) = V_{T_n} \end{cases}$$

where

$$\begin{aligned} a_n^{\text{sub}} &= \frac{1}{C} \left( g_L + g_{\text{Na}} N_m^{-3} \left( \theta^{(m)}(T_n) \right)^3 N_h^{-1} \theta^{(h)}(T_n) + g_K N_n^{-4} \left( \theta^{(n)}(T_n) \right)^4 \right) \\ b_n^{\text{sub}} &= \frac{1}{C} \left( g_L V_L + g_{\text{Na}} V_{\text{Na}} N_m^{-3} \left( \theta^{(m)}(T_n) \right)^3 N_h^{-1} \theta^{(h)}(T_n) + g_K V_K N_n^{-4} \left( \theta^{(n)}(T_n) \right)^4 \right) \\ a_n^{\text{chan}} &= \frac{1}{C} \left( g_L + g_{\text{Na}} N_{\text{Na}}^{-1} \theta^{(m_3 h_1)}(T_n) + g_K N_K^{-1} \theta^{(n_4)}(T_n) \right) \\ b_n^{\text{chan}} &= \frac{1}{C} \left( g_L V_L + g_{\text{Na}} V_{\text{Na}} N_{\text{Na}}^{-1} \theta^{(m_3 h_1)}(T_n) + g_K V_K N_K^{-1} \theta^{(n_4)}(T_n) \right) \end{aligned}$$

Then, the flow is

$$\phi_{\theta_{T_n}}(t - T_n, V_{T_n}) = e^{-a_n^x(t-T_n)} \left[ V_{T_n} + \frac{b_n^x}{a_n^x} (e^{a_n^x(t-T_n)} - 1) + \frac{1}{C} \int_{T_n}^t e^{a_n^x(s-T_n)} I(s) ds \right]$$

For both models we consider that the stimulation  $I$  takes the form  $I(t) = K \mathbf{1}_{[t_1, t_2]}(t)$  with  $K > 0$  and  $t, t_1, t_2 \in \mathbb{R}_+$ .

## 6 Simulations

We now proceed to the simulations of the *subunit model* and the *channel model* by using the thinning procedure of section 3.2. Firstly, we explicit the three bounds for both models. Secondly, we numerically compare the efficiency of the bounds in term of reject and simulation time. Finally, we use the thinning procedure to compute a variable of biological interest for both models.

### 6.1 Determination of the jump rate bounds

For simplicity of presentation, we do not distinguish the flows of the *subunit model* and the *channel model*, one has to use  $a^{\text{sub}}$  and  $b^{\text{sub}}$  for the *subunit model* and  $a^{\text{chan}}$  and  $b^{\text{chan}}$  for the *channel model*. The determination of the bounds relies on the fact that  $\alpha_n, \alpha_m, \beta_h$  are increasing functions,  $\beta_n, \beta_m, \alpha_h$  are decreasing, and that for  $n \geq 0$  the flow  $\phi(\cdot - T_n, x_{T_n})$  is bounded.

#### 6.1.1 The global bound

To determine the *global bound* we use a result in [4] concerning the *channel model* which state that if  $V_0 \in [V_-, V_+]$  then  $V_t \in [V_-, V_+] \forall t \geq 0$  with  $V_- = \min\{V_{\text{Na}}, V_K, V_L\}$  and  $V_+ = \max\{V_{\text{Na}}, V_K, V_L\}$ . By using the monotony of the opening and closing rate functions, we find

$$\tilde{\lambda}^{\text{glo}} = N_m \alpha_m(V_{\text{Na}}) + N_h \beta_h(V_{\text{Na}}) + N_n \alpha_n(V_{\text{Na}})$$

The result in [4] is also applicable to the *subunit model* and leads to the same expression of the *global bound* for this model.

### 6.1.2 The local bound

Let  $n \geq 0$  and  $t \geq T_n$ . We note  $\tilde{\lambda}_n = \sup_{t \geq T_n} \lambda(\psi(t - T_n, x_{T_n}))$ . To determine the constant  $\tilde{\lambda}_n$  we write the flow as follows

$$\phi(t - T_n, x_{T_n}) = f_n(t) + g_n(t)$$

where

$$\begin{aligned} f_n(t) &= e^{-a_n(t-T_n)} \left( V_{T_n} + \frac{b_n}{a_n} (e^{a_n(t-T_n)} - 1) \right) \\ g_n(t) &= e^{-a_n(t-T_n)} \frac{1}{C} \int_{T_n}^t e^{a_n(s-T_n)} I(s) ds \end{aligned}$$

The purpose is to determine a lower and an upper bound of the flow. We have  $a_n > 0$ ,  $b_n$  may be negative or non-negative, and  $f_n$  is monotone. By using the fact that  $\forall t \geq 0$ ,  $I(t) \leq K$  we find that

$$\underline{V}_{T_n} \leq \phi(t - T_n, x_{T_n}) \leq \bar{V}_{T_n} \quad (2)$$

where  $\bar{V}_{T_n} = V_{T_n} \vee \frac{b_n}{a_n} + \frac{K}{C a_n}$ , and  $\underline{V}_{T_n} = V_{T_n} \wedge \frac{b_n}{a_n}$ . Then, by using the monotony of the opening and closing rate functions we obtain

$$\begin{aligned} \tilde{\lambda}_n &= \left( \alpha_m(\bar{V}_{T_n})(N_m - \theta_{\text{open}}^m(T_n)) + \beta_m(\underline{V}_{T_n})\theta_{\text{open}}^m(T_n) \right) + \left( \alpha_h(\underline{V}_{T_n})(N_h - \theta_{\text{open}}^h(T_n)) + \right. \\ &\quad \left. \beta_h(\bar{V}_{T_n})\theta_{\text{open}}^h(T_n) \right) + \left( \alpha_n(\bar{V}_{T_n})(N_n - \theta_{\text{open}}^n(T_n)) + \beta_n(\underline{V}_{T_n})\theta_{\text{open}}^n(T_n) \right). \end{aligned}$$

The expression of the *local bound* is the same for the *channel* and *subunit model* but the Markov chain  $\theta$  is different.

### 6.1.3 The optimal bound

In the case of the *optimal bound*, we consider two partitions of  $[T_n, +\infty[$ . The first one is the same as in section 4.3 which is noted, for fixed  $\epsilon > 0$ ,  $(\mathcal{P}_k^{T_n, \epsilon})_{k \in \mathbb{N}}$ . We recall that, for  $k \in \mathbb{N}$ ,  $\mathcal{P}_k^{T_n, \epsilon} = [T_n + k\epsilon, T_n + (k+1)\epsilon[$ . We precise that the integrated *optimal bound* is given for  $t \geq T_n$  by

$$\tilde{\Lambda}_n^\epsilon(t) = \sum_{k \geq 0} \sup_{s \in \mathcal{P}_k^{T_n, \epsilon}} \lambda(\psi(s - T_n, x_{T_n})) \left[ (k+1)\epsilon \wedge (t - T_n) - k\epsilon \wedge (t - T_n) \right]$$

and that the inverse is given by

$$(\tilde{\Lambda}_n^\epsilon)^{-1}(t) = \sum_{p \geq 0} \left( \frac{t - \epsilon \sum_{k=0}^{p-1} \sup_{s \in \mathcal{P}_k^{T_n, \epsilon}} \lambda(\psi(s - T_n, x_{T_n}))}{\sup_{s \in \mathcal{P}_p^{T_n, \epsilon}} \lambda(\psi(s - T_n, x_{T_n}))} + T_n + p\epsilon \right) \mathbf{1}_{[\kappa_{p-1}, \kappa_p[}(t)$$

where  $\kappa_p = \epsilon \sum_{k=0}^p \sup_{s \in \mathcal{P}_k^{T_n, \epsilon}} \lambda(\psi(s - T_n, x_{T_n}))$  and, by convention,  $\kappa_{-1} = 0$ . For  $k \in \mathbb{N}$ , we have

$$\begin{aligned} \sup_{s \in \mathcal{P}_k^{T_n, \epsilon}} \lambda(\psi(s - T_n, x_{T_n})) &= \left( \alpha_m(\bar{V}_{T_n}^{k, \epsilon})(N_m - \theta_{\text{open}}^m(T_n)) + \beta_m(\underline{V}_{T_n}^{k, \epsilon})\theta_{\text{open}}^m(T_n) \right) + \\ &\quad \left( \alpha_h(\underline{V}_{T_n}^{k, \epsilon})(N_h - \theta_{\text{open}}^h(T_n)) + \beta_h(\bar{V}_{T_n}^{k, \epsilon})\theta_{\text{open}}^h(T_n) \right) + \\ &\quad \left( \alpha_n(\bar{V}_{T_n}^{k, \epsilon})(N_n - \theta_{\text{open}}^n(T_n)) + \beta_n(\underline{V}_{T_n}^{k, \epsilon})\theta_{\text{open}}^n(T_n) \right) \end{aligned}$$

Where

$$\begin{aligned}\bar{V}_{T_n}^{k,\epsilon} &= f_n(T_n + k\epsilon) \vee f_n(T_n + (k+1)\epsilon) + e^{-a_n k\epsilon} \int_{T_n}^{T_n + (k+1)\epsilon} e^{a_n(s-T_n)} I(s) ds, \\ \underline{V}_{T_n}^{k,\epsilon} &= f_n(T_n + k\epsilon) \wedge f_n(T_n + (k+1)\epsilon) + e^{-a_n(k+1)\epsilon} \int_{T_n}^{T_n + k\epsilon} e^{a_n(s-T_n)} I(s) ds.\end{aligned}$$

The second partition is obtained for  $P = \{0, 1\}$  and is noted  $(\mathcal{Q}_k^{T_n, \epsilon})_{k \in P}$  where  $\mathcal{Q}_0^{T_n, \epsilon} = [T_n, T_n + \epsilon[$  and  $\mathcal{Q}_1^{T_n, \epsilon} = [T_n + \epsilon, +\infty[$ . The integrated *optimal bound* is

$$\tilde{\Lambda}_n^\epsilon(t) = \bar{\lambda}_n^\epsilon \left[ \epsilon \wedge (t - T_n) \right] + \tilde{\lambda}_n \left[ (t - T_n) - \epsilon \wedge (t - T_n) \right]$$

Where  $\tilde{\lambda}_n$  is the previous *local bound* and

$$\bar{\lambda}_n^\epsilon = \sup_{t \in \mathcal{Q}_0^{T_n, \epsilon}} \lambda \left( \psi(t - T_n, x_{T_n}) \right).$$

The inverse is given by

$$\left( \tilde{\Lambda}_n^\epsilon \right)^{-1}(s) = \left( \frac{s}{\bar{\lambda}_n^\epsilon} + T_n \right) \mathbf{1}_{[0, \epsilon \bar{\lambda}_n^\epsilon[}(s) + \left( \frac{s - \epsilon \bar{\lambda}_n^\epsilon}{\tilde{\lambda}_n} + T_n + \epsilon \right) \mathbf{1}_{[\epsilon \bar{\lambda}_n^\epsilon, +\infty[}(s)$$

Once again, the expression of the *optimal bound* is the same for both models but the Markov chain is different. We precise that we used the *local bound* to define the *optimal bound* with the partition  $(\mathcal{Q}_k^{T_n, \epsilon})_{k \in \{0, 1\}}$ .

Note that, for  $n \geq 1$ , it is possible to define an  $\epsilon_n$  which is "adapted" to the inter jump time  $T_{n+1} - T_n$ . To determine such an  $\epsilon_n$ , we use the bounds of the flow in inequality (2) to define a lower *local bound*  $\underline{\lambda}_n$  such that

$$\begin{aligned}\underline{\lambda}_n &= \left( \alpha_m(\underline{V}_{T_n})(N_m - \theta_{\text{open}}^m(T_n)) + \beta_m(\bar{V}_{T_n})\theta_{\text{open}}^m(T_n) \right) + \left( \alpha_h(\bar{V}_{T_n})(N_h - \theta_{\text{open}}^h(T_n)) + \right. \\ &\quad \left. \beta_h(\underline{V}_{T_n})\theta_{\text{open}}^h(T_n) \right) + \left( \alpha_n(\underline{V}_{T_n})(N_n - \theta_{\text{open}}^n(T_n)) + \beta_n(\bar{V}_{T_n})\theta_{\text{open}}^n(T_n) \right).\end{aligned}$$

We note  $(\underline{T}_n)_{n \geq 0}$  the corresponding point process, then we have

$$0.05 = \mathbb{P}(T_{n+1} - T_n > \epsilon_n) \leq \mathbb{P}(\underline{T}_{n+1} - \underline{T}_n > \epsilon_n) = e^{-\epsilon_n \underline{\lambda}_n}$$

and we take

$$\epsilon_n = \frac{-\log(0.05)}{\underline{\lambda}_n}.$$

Note that the value  $\epsilon_n$  is in fact adapted to the inter jump time  $\underline{T}_{n+1} - \underline{T}_n$ .

## 6.2 Numerical results

In this section, we numerically compare the three different jump rate bounds and we use thinning algorithm to simulate a variable of biological interest, the spiking times.

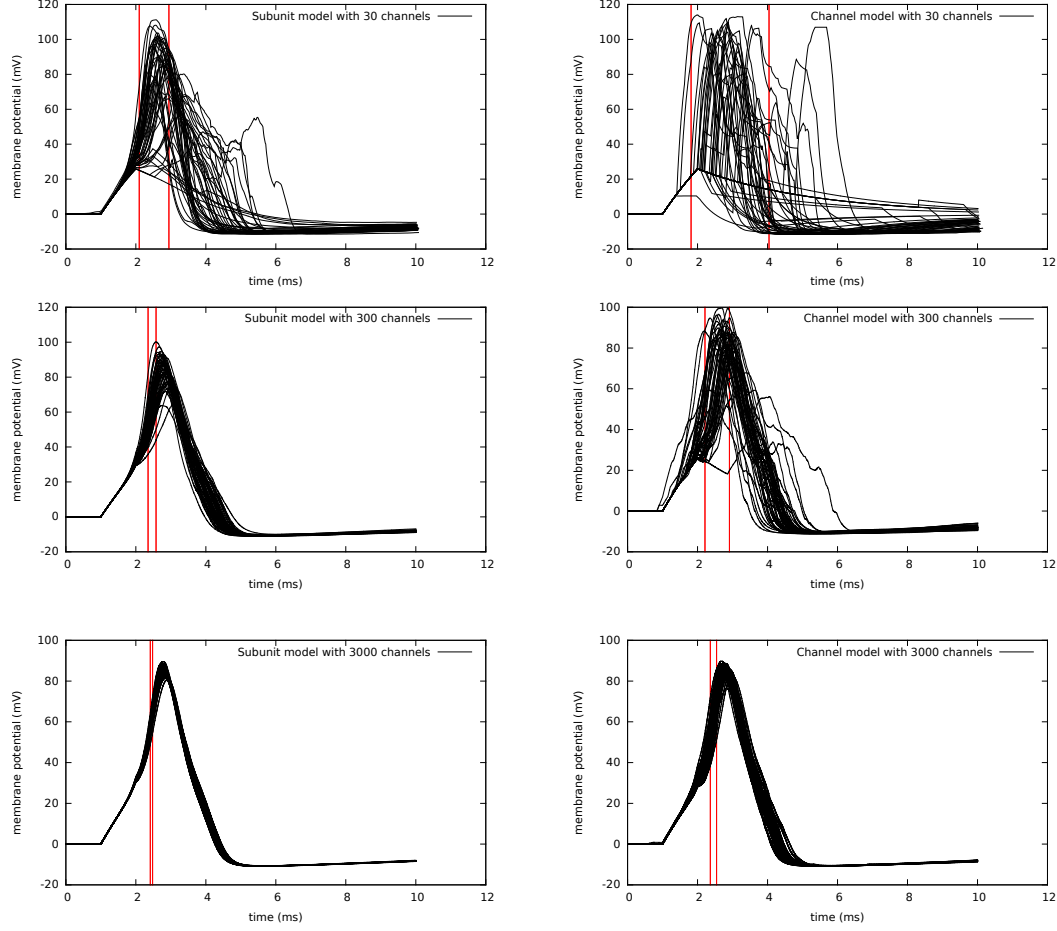


Figure 2: First column : *subunit model*. Second column : *channel model*. Vertical lines are the standard deviation of the spiking times (see section 6.2.2).

### 6.2.1 Numerical comparison of the jump rate bounds

In this part, we first show trajectories of the two stochastic Hodgkin-Huxley models obtained with the thinning method using the *optimal bound*. Then, we collect in several tables and graphs the results concerning the simulation time and the rate of acceptance of both models.

In the sequel, for  $\epsilon > 0$ , the *optimal- $\mathcal{Q}^\epsilon$*  (respectively *optimal- $\mathcal{P}^\epsilon$* ) bound denotes the optimal bound using the partition  $(\mathcal{Q}_k^{T_n, \epsilon})_{k \in \{0,1\}}$  (respectively  $(\mathcal{P}_k^{T_n, \epsilon})_{k \in \mathbb{N}}$ ). All numerical values are obtained from a classical Monte Carlo method with 100 000 trials. Parameters of the models are given in Appendix A - B. We denote by  $N_{\text{chan}}$  the common number of sodium and potassium channels,  $N_{\text{chan}} = N_{\text{Na}} = N_{\text{K}}$ . The input current is  $I(t) = 30\mathbf{1}_{[1,2]}(t)$ . The simulation time represents the time needed to simulate one path of the PDMP on  $[0,10]$ .

Each lines of Figure 2 shows fifty trajectories of the *subunit* and the *channel model* with a different number of channels,  $N_{\text{chan}} = 30, 300, 3000$ . It allows to see the different behaviours of the two models. In each lines, we see that the behaviour of the *channel model* is more erratic than the *subunit model* one (except for the third line where the two models have approximately

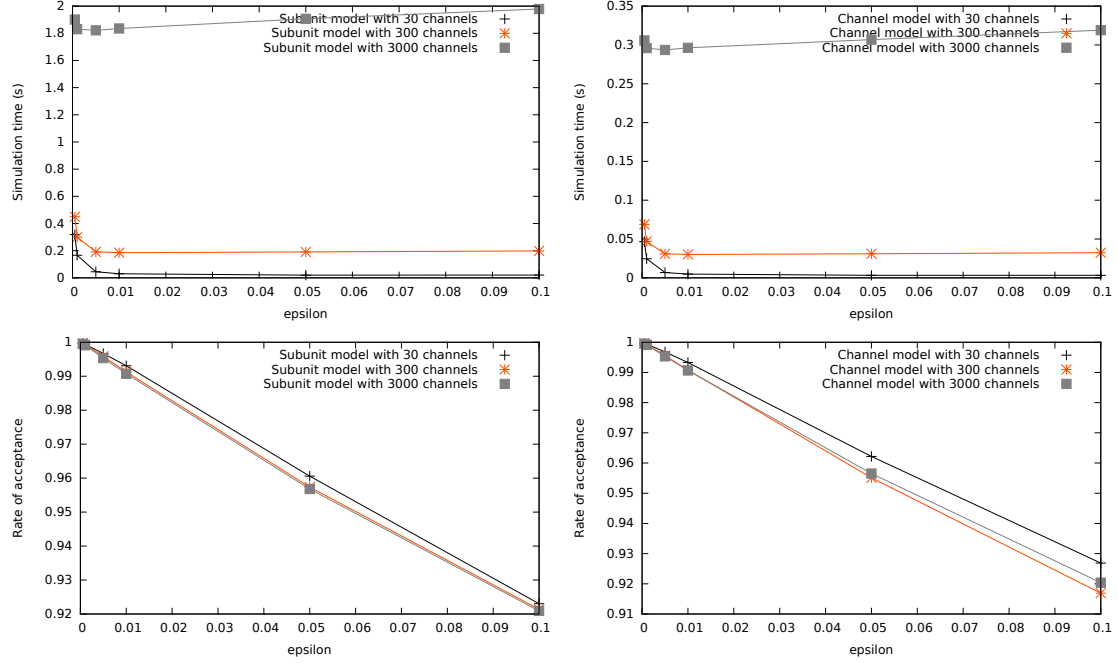


Figure 3: Simulation time and rate of acceptance with the optimal- $\mathcal{P}^\epsilon$  bound as a function of the parameter  $\epsilon$ .

the same behaviour). Differences in trajectories are mainly explained by two distinct modelling approaches of the conductance of the membrane. In the *subunit model*, we consider that the conductance at time  $t$  depends on the fraction of open gates at time  $t$ , thus, the equation of the voltage changes rapidly at the same time as the state of the gates. In the *channel model*, the conductance at time  $t$  depends on the fraction of active channels at time  $t$ , therefore, a change in the state of the gates may not implies a change in the voltage's equation. Thus, the dynamic of the membrane potential change less than in the first case and trajectories are more irregular. We also see that more the number of channels is high more the differences in trajectories are small. It illustrates a result in [24] where the authors showed that the deterministic limit when the number of channels goes to infinity of both models are equivalent. However, it seems that the convergence speed is not the same.

Concerning the *optimal- $\mathcal{P}^\epsilon$*  bound, we see on Figure 3 that in both models, more  $\epsilon$  is small less rejected points are. It illustrates the fact that  $\tilde{N}^\epsilon$  converge to  $N$  when  $\epsilon$  goes to 0 (proposition 4.4). Figure 3 also shows that, for fixed  $N_{\text{chan}}$ , the simulation time varies with  $\epsilon$ . For both models, the value of  $\epsilon$  which minimize the simulation time is inversely proportional to the parameter  $N_{\text{chan}}$ . Let  $\epsilon(N_{\text{chan}})$  be that optimal value of  $\epsilon$ . For increasing  $\epsilon > \epsilon(N_{\text{chan}})$ , the rate of acceptance decreases, thus, we have to simulate more and more uniform pseudo-random variables and the simulation time increases. For decreasing  $\epsilon < \epsilon(N_{\text{chan}})$ , the rate of acceptance increases but the simulation time too because of the increasing number of iterations needed to compute the integrated jump rate bound and its inverse. Thus, one has to take a small (respectively large)  $\epsilon$  when the jumps frequency is high (respectively low).

We see on Figure 4 that more  $\epsilon$  is small more the rate of acceptance of the *optimal- $\mathcal{Q}^\epsilon$*  bound is close to the one of the *local bound*. Note that the value of  $\epsilon$  which maximise the rate of acceptance



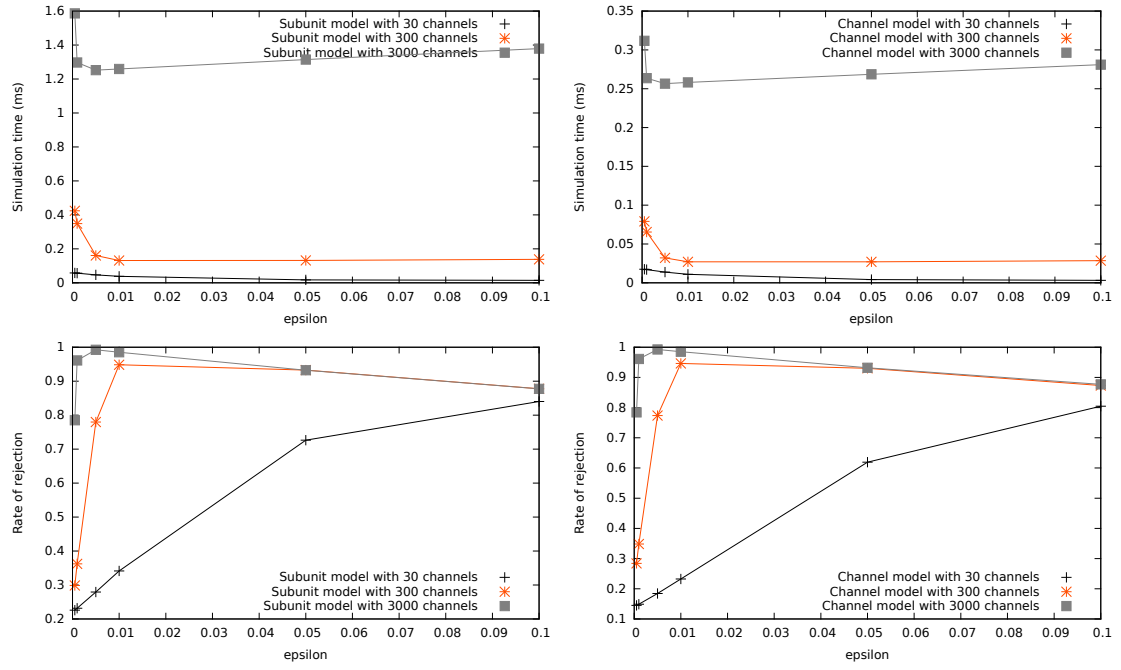


Figure 4: Simulation time and rate of acceptance with the optimal- $Q^\epsilon$  bound as a function of the parameter  $\epsilon$ .

Table 1: Simulation time and rate of acceptance for  $N_{\text{chan}} = 30$ . The lines ODE represent the algorithm in [25] with  $h = 10^{-3}$  for both *subunit model* and *channel model*

Model	Bound	simulation time (sec)	rate of acceptance
Channel	Optimal- $\mathcal{Q}^{\epsilon_n}$	0,003 ( $\pm 8.10^{-7}$ )	0,857 ( $\pm 2.10^{-3}$ )
	Local	0,008 ( $\pm 6.10^{-6}$ )	0,141 ( $\pm 2.10^{-3}$ )
	Global	0,012 ( $\pm 3.10^{-6}$ )	0,065 ( $\pm 6.10^{-5}$ )
	ODE	0.009 ( $\pm 1.10^{-7}$ )	
Subunit	Optimal- $\mathcal{Q}^{\epsilon_n}$	0,016 ( $\pm 1.10^{-6}$ )	0,88 ( $\pm 1.10^{-3}$ )
	Local	0,050 ( $\pm 2.10^{-4}$ )	0,22 ( $\pm 1.10^{-3}$ )
	Global	0,12 ( $\pm 3.10^{-4}$ )	0,061 ( $\pm 2.10^{-5}$ )
	ODE	0.016 ( $\pm 2.10^{-7}$ )	

Table 2: Simulation time and rate of acceptance for  $N_{\text{chan}} = 300$ . The lines ODE represent the algorithm in [25] with  $h = 10^{-4}$  for both *subunit model* and *channel model*.

Model	Bound	simulation time (sec)	rate of acceptance
Channel	Optimal- $\mathcal{Q}^{\epsilon_n}$	0,030 ( $\pm 3.10^{-5}$ )	0,962 ( $\pm 9.10^{-5}$ )
	Local	0,050 ( $\pm 1.10^{-4}$ )	0,223 ( $\pm 3.10^{-4}$ )
	Global	0,120 ( $\pm 3.10^{-4}$ )	0,062 ( $\pm 7.10^{-5}$ )
	ODE	0.094 ( $\pm 1.10^{-5}$ )	
Subunit	Optimal- $\mathcal{Q}^{\epsilon_n}$	0,148 ( $\pm 5.10^{-4}$ )	0,957 ( $\pm 9.10^{-5}$ )
	Local	0,244 ( $\pm 1.10^{-3}$ )	0,237 ( $\pm 8.10^{-5}$ )
	Global	0,322 ( $\pm 2.10^{-3}$ )	0,061 ( $\pm 1.10^{-5}$ )
	ODE	0.157 ( $\pm 1.10^{-5}$ )	

is the same which minimize the simulation time. As in the case of the *optimal- $\mathcal{P}^\epsilon$*  bound, the optimal value of  $\epsilon$  is inversely proportional to  $N_{\text{chan}}$ . For decreasing  $\epsilon < \epsilon(N_{\text{chan}})$ , the rate of acceptance decreases and the simulation time increases because we use the *local bound*  $\tilde{\lambda}_n$  instead of the smaller bound  $\bar{\lambda}_n^\epsilon$  (see section 6.1.3). For increasing  $\epsilon > \epsilon(N_{\text{chan}})$ , the rate of acceptance decreases and the simulation time increases because the bound  $\bar{\lambda}_n^\epsilon$  becomes bigger and bigger. By comparing the *optimal- $\mathcal{Q}^{\epsilon(N_{\text{chan}})}$*  and the *optimal- $\mathcal{P}^{\epsilon(N_{\text{chan}})}$*  bound we see that the first one is the most efficient in term of simulation time, it is also the simplest to implement. However, this bound does not exist when the jump rate or the flow is not bounded. In this case, one may use the *optimal- $\mathcal{P}^{\epsilon(N_{\text{chan}})}$*  bound which is efficient too but more complex to implement.

From figure 3 and 4, we see that for both the *optimal- $\mathcal{Q}^\epsilon$*  and the *optimal- $\mathcal{P}^\epsilon$*  bound the best simulation time is achieved for  $\epsilon(30) = 0.1$ ,  $\epsilon(300) = 0.01$  and  $\epsilon(3000) = 0.005$ . We saw in sections 5.2.1 and 5.2.2 that the *subunit model* and the *channel model* share the same jump rate. For both models, the maximum value of the inter-jump times is of order  $10^{-1}$  for  $N_{\text{chan}} = 30$ ,  $10^{-2}$  for  $N_{\text{chan}} = 300$  and  $10^{-3}$  for  $N_{\text{chan}} = 3000$ . It coincides with the values  $\epsilon(N_{\text{chan}})$  which, in this case, confirm that the optimal simulation time is obtained for  $\epsilon$  of order  $\max_n |T_{n+1} - T_n|$ .

Tables 1-3 show results of numerical computations of the simulation time and of the rate

Table 3: Simulation time and rate of acceptance for  $N_{\text{chan}} = 3000$ . The lines ODE represent the algorithm in [25] with  $h = 10^{-5}$  for both *subunit model* and *channel model*.

Model	Bound	simulation time (sec)	rate of acceptance
Channel	optimal- $\mathcal{Q}^{\epsilon_n}$	0,296 ( $\pm 3.10^{-3}$ )	0,965 ( $\pm 2.10^{-5}$ )
	Local	0,474 ( $\pm 6.10^{-3}$ )	0,236 ( $\pm 3.10^{-5}$ )
	Global	1,184 ( $\pm 2.10^{-2}$ )	0,060 ( $\pm 3.10^{-7}$ )
	ODE	0.940 ( $\pm 5.10^{-4}$ )	
Subunit	Optimal- $\mathcal{Q}^{\epsilon_n}$	1,471 ( $\pm 3.10^{-2}$ )	0,964 ( $\pm 9.10^{-6}$ )
	Local	2,478 ( $\pm 4.10^{-2}$ )	0,238 ( $\pm 7.10^{-6}$ )
	Global	3,315 ( $\pm 3.10^{-1}$ )	0,060 ( $\pm 9.10^{-8}$ )
	ODE	1.567 ( $\pm 1.10^{-3}$ )	

of acceptance of the thinning method for the *global*, *local* and *optimal- $\mathcal{Q}^{\epsilon_n}$*  bound using both the *channel* and the *subunit model* with different values of the parameter  $N_{\text{chan}}$ . For both models and for all the studied values of  $N_{\text{chan}}$ , the simulation time using the *optimal* bounds ( $\mathcal{Q}^{\epsilon(N_{\text{chan}})}, \mathcal{P}^{\epsilon(N_{\text{chan}})}$  and  $\mathcal{Q}^{\epsilon_n}$ ) is better than the one obtained with both the *global* and *local bound*. Note that the *optimal- $\mathcal{Q}^{\epsilon_n}$*  bound is more efficient than the *optimal- $\mathcal{P}^{\epsilon(N_{\text{chan}})}$*  bound to simulate the *subunit model*. Since the computation of  $\epsilon_n$  requires the computation of the jump rate bound at each iterations, the *optimal- $\mathcal{Q}^{\epsilon_n}$*  bound will be more efficient when the jumps frequency is low. Thus, for all studied values of  $N_{\text{chan}}$ , the *optimal- $\mathcal{Q}^{\epsilon(N_{\text{chan}})}$*  is the most efficient. The differences of simulation time between the *subunit* and the *channel model* are explained by the fact that the numerical computation of the flow of the *channel model* is cheaper than the one of the *subunit model*. Note that the simulation time using the three bounds (*global*, *local*, *optimal*) increases linearly as a function of  $N_{\text{chan}}$ .

In the ODE algorithm [25], we need to adapt the time step  $h$  when the parameter  $N_{\text{chan}}$  varies, otherwise, we do not simulate the expected trajectories. Thinning algorithm in the *channel model* speeds up the simulation by a factor 3 compared to the ODE method whereas in the *subunit model* the factor is approximately 1. Such a difference is explained by the fact that the ratio of the simulation times between the flows of the *subunit* and the *channel* (for thinning algorithm) is bigger than the ratio of the simulation times between the vector fields of the *subunit* and the *channel* (for ODE algorithm).

Despite the complexity of the *optimal bound* compared to the two others, it is the most efficient one to simulate both the *channel model* and the *subunit model*.

### 6.2.2 Spiking times

Bio-scientists believe that the timing of action potentials is one of the characteristics of the nervous system which carries the most of informations. It has been shown experimentally [29] that if a neuron is repeatedly stimulated by identical pulses, both the amplitude and the timing of the action potentials is variable. In the sequel we numerically compare the mean value of the spiking time of the *subunit* and *channel model* to the one of the deterministic Hodgkin-Huxley model.

Let  $(x_t)$  be the *subunit model* or the *channel model* defined on a filtered probability space  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}_x)$ . We consider that the stimulation is a monophasic current which produces only

one action potential within a given time window  $[0, T]$  as in Figure 1. We suppose that a spike occurs when the membrane potential exceeds a certain value noted  $\nu$ . Let  $\mathcal{T}$  be the spiking time defined by

$$\mathcal{T} = \inf\{t \in [0, T] : V_t \geq \nu\}$$

We are interested in the numerical computation of the mean and the standard deviation of  $\mathcal{T}$  as a function of the number of channels. For low values of the parameters  $N_{Na}$  and  $N_K$  a spike may never occur. In this case,  $\mathcal{T} = T$  and we do not count these trajectories in the Monte Carlo procedure. Thus, we evaluate the mean value of the spiking time conditionally on having a spike,  $\mathbb{E}[\mathcal{T} | \mathcal{T} < T]$ , with the following estimator  $I_M = (1/M) \sum_{k=1}^M \mathcal{T}_k$  where  $(\mathcal{T}_k)$  are iid realizations of  $\mathcal{T}$  conditionally on  $\{\mathcal{T} < T\}$ .

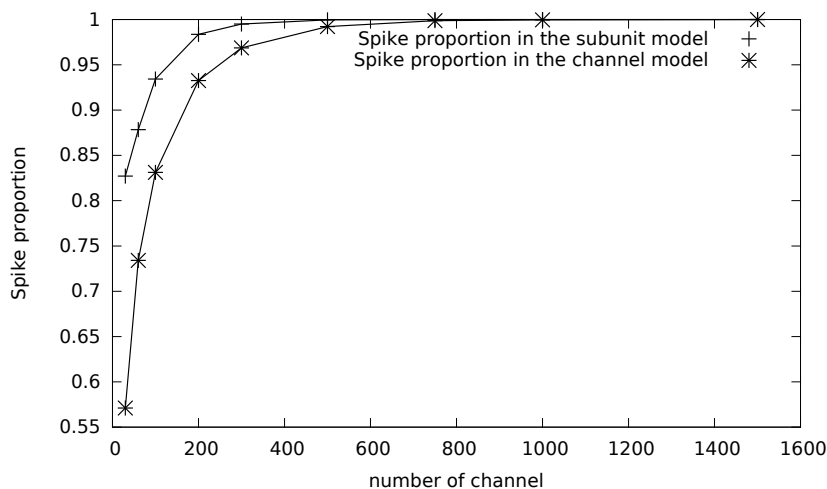


Figure 5: The graph represents the proportion of spikes obtained with the *subunit model* and the *channel model* as a function of the number of channels  $N_{\text{chan}}$ .

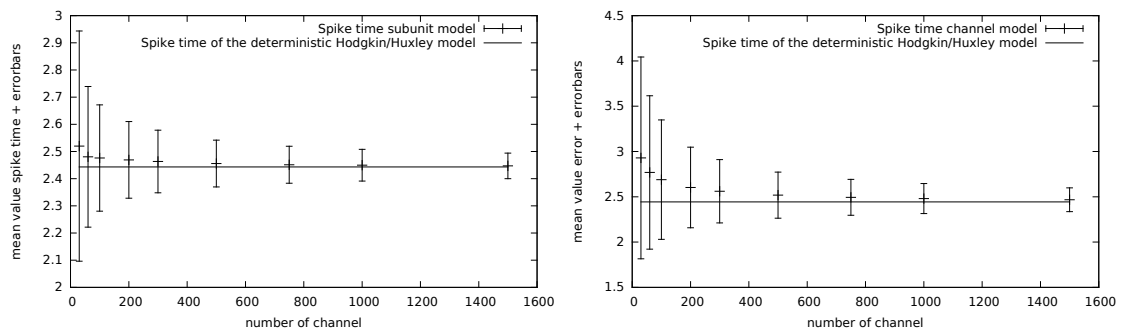


Figure 6: Mean value of the spiking time (ms) with standard deviation as a function of the number of channels  $N_{\text{chan}}$ . Left: *subunit model*. Right: *channel model*.

It has been shown in [24] that the deterministic limits of both the *subunit* (Hodgkin-Huxley of dimension four [16]) and the *channel model* (Hodgkin-Huxley of dimension fourteen [24]) are equivalent when the initial conditions satisfy a combinatorial relationship. We consider that, at

time  $t = 0$ , all the gates of the *subunit model* are closed and all the channels of the *channel model* are in the corresponding state, i.e state  $\{m_0h_0\}$  for the sodium and  $\{n_0\}$  for the potassium. These initial conditions satisfy the combinatorial relationship in [24]. The initial conditions of both deterministic Hodgkin-Huxley models are also chosen so that they satisfy the binomial relation. Thus, the spiking time of these deterministic models is the same. In the simulations, we take  $T = 10$ ,  $\nu = 60$ , we consider that the stimulation is given by  $I(t) = 301\mathbf{1}_{[1,2]}(t)$  and that  $N_{\text{Na}} = N_{\text{K}} = N_{\text{chan}}$ . In this case, the spiking time of the deterministic model is  $\mathcal{T}^{\text{deter}} = 2,443$ . Figure 6 illustrates the convergence of the mean spiking time of both the *subunit* and the *channel model* when the number of channels goes to infinity. For  $N_{\text{chan}} = 1500$  we see that the dispersion of the spiking time around its deterministic limit is approximately of order  $10^{-1}$  ms for the *subunit model* and of order  $10^{-2}$  ms for the *channel model*. Thus, a membrane patch with a number of channels superior to 1500 mimics the behaviour of the deterministic Hodgkin-Huxley model. For a number of channels inferior to 500, we see from figure 5 that the neuron may not respond to the stimuli. In this case, the dispersion of the spiking time ranges from approximately  $10^{-1}$  and almost 1 ms and are consistent with the observations in [29]. Since the simulation is exact the estimator  $I_M$  is unbiased and errors due to the Monte Carlo procedure are of order of  $M^{-1/2}$ .

## A Rate functions and parameters

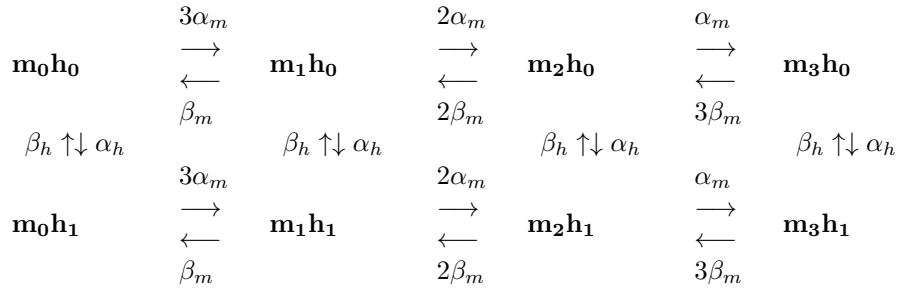
$$\alpha_n(V) = \frac{(0.1-0.01V)}{\exp(1-0.1V)-1}, \quad \alpha_m(V) = \frac{(2.5-0.1V)}{\exp(2.5-0.1V)-1}, \quad \alpha_h(V) = 0.07 \exp(-\frac{V}{20}),$$

$$\beta_n(V) = 0.125 \exp(-\frac{V}{80}), \quad \beta_m(V) = 4 \exp(-\frac{V}{18}), \quad \beta_h(V) = \frac{1}{\exp(3-0.1V)+1},$$

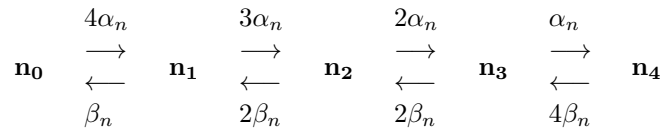
$$V_{\text{Na}} = 115, \quad g_{\text{Na}} = 120, \quad V_{\text{K}} = -12, \quad g_{\text{K}} = 36, \quad V_{\text{L}} = 0, \quad g_{\text{L}} = 0.3, \quad C = 1.$$

## B Markov schemes for the channel model

Sodium (Na) scheme:



Potassium (K) scheme:



## C Rate of acceptance for the thinning of Poisson processes

Let  $N$  and  $\tilde{N}$  be two Poisson processes with jump rate  $\lambda$  and  $\tilde{\lambda}$  respectively and jump times  $(T_n)_{n \geq 1}$  and  $(\tilde{T}_n)_{n \geq 1}$  respectively. Assume that  $N$  is the thinning of  $\tilde{N}$ . Since  $\mathbb{P}(\tilde{N}_t = 0) = e^{-\int_0^t \tilde{\lambda}(s) ds}$ , we define the rate of acceptance by  $\mathbb{E}[N_t/\tilde{N}_t | \tilde{N}_t \geq 1]$ . In the case of Poisson processes this indicator takes the following form

$$\mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] = \frac{\int_0^t \lambda(s) ds}{\int_0^t \tilde{\lambda}(s) ds} \quad (7)$$

To get (7), we use the following result which is similar to the  $n$ -uplet of non-ordering uniform variables in the Poisson homogeneous case

$$f_{(\tilde{T}_1, \dots, \tilde{T}_n | \tilde{N}_t = n)}(t_1, \dots, t_n) = \frac{\tilde{\lambda}(t_1) \dots \tilde{\lambda}(t_n)}{\left(\int_0^t \tilde{\lambda}(s) ds\right)^n} \mathbf{1}_{(t_1, \dots, t_n) \in [0, t]^n} \quad (8)$$

Equation (8) gives an explicit formula of the conditional density of the vector  $(\tilde{T}_1, \dots, \tilde{T}_n | \tilde{N}_t = n)$ . Note that we do not consider any ordering in points  $(\tilde{T}_k)_{0 \leq k \leq n}$  and that conditionally to  $\tilde{N}_t = n$ , the points  $\tilde{T}_1, \dots, \tilde{T}_n$  are independent with density  $\left(\tilde{\lambda}(s) / \int_0^t \tilde{\lambda}(u) du\right) \mathbf{1}_{s \in [0, t]}$ .

With (8) one is able to determine that

$$\mathcal{L}(N_t | \tilde{N}_t = n) = \mathcal{B}(n, p) \quad (9)$$

with  $p = \int_0^t \lambda(s) ds / \int_0^t \tilde{\lambda}(s) ds$  by noting that for  $k \leq n$

$$\{N_t = k | \tilde{N}_t = n\} = \bigcup_{1 \leq i_1 < \dots < i_k \leq n} \left[ \bigcap_{i \in \{i_1, \dots, i_k\}} \{U_i \leq \frac{\lambda}{\tilde{\lambda}}(\tilde{T}_i) | \tilde{N}_t = n\} \bigcap_{i \in \{i_1, \dots, i_k\}^c} \{U_i > \frac{\lambda}{\tilde{\lambda}}(\tilde{T}_i) | \tilde{N}_t = n\} \right]$$

and then

$$\mathbb{P}(N_t = k | \tilde{N}_t = n) = \binom{n}{k} \mathbb{P}\left(U_i \leq \frac{\lambda}{\tilde{\lambda}}(\tilde{T}_i) | \tilde{N}_t = n\right)^k \mathbb{P}\left(U_i > \frac{\lambda}{\tilde{\lambda}}(\tilde{T}_i) | \tilde{N}_t = n\right)^{n-k}$$

where  $(U_i)$  are independent variables uniformly distributed in  $[0, 1]$ , independent of  $(\tilde{T}_i)$ . Thus, the law of the number of accepted points is binomial conditionally on the number of proposed points. Then, we find (7) by noting that

$$\mathbb{E}\left[\frac{N_t}{\tilde{N}_t} | \tilde{N}_t \geq 1\right] = \frac{1}{\mathbb{P}(\tilde{N}_t \geq 1)} \sum_{n \geq 1} \frac{1}{n} \mathbb{E}[N_t | \tilde{N}_t = n] \mathbb{P}(\tilde{N}_t = n)$$

We should note that to find the explicit formula (3), we work almost exclusively on the process  $(\tilde{N}_t)$ . The only characteristic of  $N$  that we use is its jump rate  $\lambda$  and we use it to evaluate the thinning probabilities  $(\lambda/\tilde{\lambda})$ .

## References

- [1] A. Alfonsi, E. Cancés, G. Turinci, B. Di Ventura, and W. Huisinga. Adaptive hybrid simulation of hybrid stochastic and deterministic models for biochemical reactions. *ESAIM: proceedings*, 14:1–13, 2005.

- [2] D.F. Anderson, B. Ermentrout, and P.J. Thomas. Stochastic representation of ion channel kinetics and exact stochastic simulation of neuronal dynamics. *Journal of Computational Neuroscience*, 38:67–82, 2015.
- [3] F. Bouguet. Quantitative speeds of convergence for exposure to food contaminants. *ESAIM: Probability and Statistics*, 19:482–501, 2015.
- [4] E. Buckwar and M.G. Riedler. An exact stochastic hybrid model of excitable membranes including spatio-temporal evolution. *Journal of Mathematical Biology*, 63:1051–1093, 2011.
- [5] D. Chafaï, F. Malrieu, and K. Paroux. On the long time behavior of the tcp window size process. *Stochastic Processes and their Applications*, 120:1518–1534, 2010.
- [6] C.C. Chow and J.A. White. Spontaneous action potentials due to channel fluctuations. *Biophysical Journal*, 71:3013–3021, 1996.
- [7] J.R. Clay and DeFelice. Relationship between membrane excitability and single channel open-close kinetics. *Biophysical Journal*, 42:151–157, 1983.
- [8] C. Coccozza-Thivent. *Processus stochastiques et fiabilité des systèmes*. Springer, 1997.
- [9] M.H.A. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal statistical Society*, 46:353–388, 1984.
- [10] M.H.A. Davis. *Markov Models and Optimization*. Chapman and Hall, London, 1993.
- [11] L. Devroye. *Non-uniform random variate generation*. Springer-Verlag, New York Inc., 1986.
- [12] M. Doumic, M. Hoffmann, N. Krell, and L. Robert. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli Society for Mathematical Statistics and Probability*, 21:1760–1799, 2015.
- [13] R.F. Fox. Stochastic Versions of the Hodgkin-Huxley Equations. *Biophysical Journal*, 72:2068–2074, 1997.
- [14] J.H. Goldwyn, N.S. Imennov, M. Famular, and E. Shea-Brown. Stochastic differential equation models for ion channels noise in Hodgkin-Huxley neurons. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 83 (4 Pt 1), 2011.
- [15] B. Hille. *Ionic channels of excitable membranes*. Sinauer Associates Inc, 1992.
- [16] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.
- [17] O. Kallenberg. *Random measures*. Akademie-Verlag Berlin, 1975.
- [18] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, 2002.
- [19] P.A.W. Lewis and G.S. Shedler. Simulation of nonhomogeneous Poisson processes by thinning. *Naval Research Logistics Quarterly*, 26:403–413, 1979.
- [20] H. Mino, J.T. Rubinstein, and J.A White. Comparison of algorithms for the simulation of action potentials with stochastic sodium channels. *Annals of Biomedical Engineering*, 30:578–587, 2002.

- [21] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical Journal*, 35:193–213, 1981.
- [22] Y. Ogata. On Lewis’ Simulation Method for Point Processes. *IEEE Transactions on Information Theory*, 27, 1981.
- [23] P. Orio and D. Soudry. Simple, fast and accurate implementation of the diffusion approximation algorithm for stochastic ion channels with multiple states. *PLoS one*, 7, 2012.
- [24] K. Pakdaman, M. Thieullen, and G. Wainrib. Fluid limit theorems for stochastic hybrid systems with application to neuron models. *advanced applied probability*, 42:761–794, 2010.
- [25] M.G. Riedler. Almost sure convergence of numerical approximations for Piecewise Deterministic Markov Processes. *Journal of Computational and Applied Mathematics*, pages 50–71, 2012.
- [26] J.T. Rubinstein. Threshold fluctuations in an N sodium channel model of the node of ranvier. *Biophysical Journal*, 68:779–785, 1995.
- [27] E. Skaugen and L. Walloe. Firing behaviour in a stochastic nerve membrane model based upon the Hodgkin-Huxley equations. *Acta Physiologica*, 107:343–363, 1979.
- [28] R. Veltz. A new twist for the simulation of hybrid systems using the true jump method. *arXiv[math]*, 2015.
- [29] A.A. Verveen and H.E. Derksen. Fluctuation phenomena in nerve mebrane. *proceedings of the IEEE*, 56:906–916, 1968.